

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 06:53:49 ; Search time 1010.34 Seconds
(without alignments)
8030.908 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300

Perfect score: 501

Sequence: 1 caattgtggaagcttgcgg.....ttgtcttaacaagtggt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estrom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489	97.6	835	9	AF188523
2	483	96.4	695	9	AL133724
3	483	96.4	777	13	BI602169
4	479.8	95.6	644	10	AW162535
5	479	95.6	920	12	BG676871
6	437.2	87.3	801	13	BI603180

Result No.	Score	Query Match	Length	DB ID	Description
7	427.2	85.3	546	12	BF748808
8	410.8	82.0	679	10	AW953240
9	385.2	76.9	574	12	BF667219
10	369	73.7	521	13	BM353207
11	368	73.5	539	13	BM353362
12	359	71.7	359	9	AA484945
13	311.4	62.2	609	13	BG992490
14	301	60.1	870	12	BG425739
15	299	58.6	657	9	AI372818
16	293.8	58.6	657	10	BB309170
17	282.8	56.4	359	9	AA705585
18	272.6	54.4	371	9	AA151055
19	271	54.1	310	9	AI372817
20	269.8	53.9	483	12	BG384511
21	263.8	52.7	523	10	BE372746
22	254	50.7	896	12	BG468527
23	251	50.1	501	12	BF044659
24	250	49.9	658	13	BI153873
25	239.6	47.8	374	10	AA484613
26	237.8	47.5	371	9	AU083157
27	233	46.5	388	13	BI050739
28	225.6	45.0	377	12	BF042085
29	217	43.3	491	14	BQ644790
30	216.4	43.2	594	10	AV720937
31	207.6	41.4	765	9	AL553184
32	207.2	41.4	328	9	AA879884
33	206.8	41.3	394	13	BI337284
34	205	40.9	749	12	BF695548
35	203.8	40.7	450	12	BE703924
36	202.2	40.4	352	14	T03019
37	192.2	38.4	334	12	BF041935
38	188	37.5	1039	9	AL570666
39	158.4	31.6	333	10	AA637577
40	156.2	31.2	370	12	BE930535
41	149.2	29.8	321	10	AM227478
42	146.4	29.2	535	12	BF963798
43	145.4	29.0	184	12	BE699344
44	143	28.5	512	12	BF090072
45	141.6	28.3	705	13	BG922853

ALIGNMENTS

RESULT 1
AF188523/C
LOCUS
DEFINITION
AF188523 Homo sapiens ATCC HTB-12; SW1088 Homo sapiens CDNA clone
ISG 4, mRNA sequence.
ACCESSION
AF188523
VERSION
AF188523.1 GI:7144569
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
Ye, Z. and Connor, J. R.
TITLE
Identification of Iron Regulated Genes by Rescreening CDNA
Libraries from SSH with Antisense Probe from Three Iron Conditions
JOURNAL
Unpublished (2000)
COMMENT
Contact: Ye Z
Neuroscience and Anatomy
Pennsylvania State University College of Medicine
500 University Drive, Hershey, PA 17033, USA
Library screened by SSH and Reverse Northern blot; increased
expression in iron loading was confirmed by Northern blot.
FEATURES
source
1..835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ISG 4"
/clone_lib="Homo sapiens ATCC HTB-12; SW1088"

/issue_type="astrocytoma"
/cell_line="ATCC HTB-12; SW1088"
/note="Organ: liver"

BASE COUNT 225 a 209 c 156 g 245 t

Query Match 97.6%; Score 489; DB 9; Length 835;
Best Local Similarity 99.8%; Pred. No. 3.4e-117;
Matches 500; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CAATTTGTGGAACCTTCCGCTCAAGCCTTAGCGGTAAAGAACGCTGTGATTAAAGAGAC 60
DB 769 CAATTTGTGGAACCTTCCGCTCAAGCCTTAGCGGTAAAGAACGCTGTGATTAAAGAGAC 710
QY 61 CAGCTGAGATATGAGAGAAATGAAGCAACTACAGGGAATGGCGAAGAGCTTTCT 120
DB 709 CAGCTGAGATATGAGAGAAATGAAGCAACTACAGGGAATGGCGAAGAGCTTTCT 650
QY 121 GAAATCATGCATGACAGATCTGCCCCCTGGAGAGAGACGAGCTCTTACGGAATTC 180
DB 649 GAAATCATGCATGACAGATCTGCCCCCTGGAGAGAGACGAGCTCTTACGGAATTC 590
QY 181 CTTCACATCTTCAAGCCATCACTGAGGACTCCACAAGACCAATGTTACGGGATGACC 240
DB 589 CTTCACATCTTCAAGCCATCACTGAGGACTCCACAAGACCAATGTTACGGGATGACC 530
QY 241 AGCTCGTCTCGTCTGATGATTACATCTCATGCCCCGTGTGGGGACTTCTTGTC 300
DB 529 AGCTCGTCTCGTCTGATGATTACATCTCATGCCCCGTGTGGGGACTTCTTGTC 470
QY 301 TTTCGCAACTCAGATGCTTTCCAAAGCCATCACTGAGGAGACGAGACAGGAGAGAC 360
DB 469 TTTCGCAACTCAGATGCTTTCCAAAGCCATCACTGAGGAGACGAGACAGGAGAGAC 410
QY 361 CAAGGGAGAGGGAGAGAAAGAAATGAAGCAACAGTATTCTTAAAGACTTTCTAT 420
DB 409 CAAGGGAGAGGGAGAGAAAGAAATGAAGCAACAGTATTCTTAAAGACTTTCTAT 350
QY 421 AGGAGTTGTAGAGAGTGCACATATTTTAAATCTCACTGCAATATTCGAAATTTTC 480
DB 349 AGGAGTTGTAGAGAGTGCACATA-TTTTAAATCTCACTGCAATATTCGAAATTTTC 291
QY 481 ATTGTCTTAAACAAAGTGT 501
DB 290 ATTGTCTTAAACAAAGTGT 270

RESULT 2
AL133724 695 bp mRNA linear EST 25-FEB-2000

LOCUS DKEP761J19121_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
DEFINITION DKEP761J19121 5', mRNA sequence.

ACCESSION AL133724
VERSION AL133724.1 GI:6601912
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 695)
Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Poustka, et al.)
Unpublished (1999)
Contact: Poustka A.J.
Department Lehrbuch
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No SI sequence available.
This clone (DKEP761J19121) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 695

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKEP761J19121"
/clone_1id="761 (synonym: hamy2)"
/issue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSPori1, Site_1: NotI; Site_2: SalI"

BASE COUNT 202 a 134 c 185 g 174 t

Query Match 96.4%; Score 483; DB 9; Length 695;
Best Local Similarity 98.4%; Pred. No. 1.2e-115;
Matches 501; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 CAATTTGTGGAACCTTCCGCTCAAGCCTTAGCGGTAAAGAACGCTGTGATTAAAGAGAC 60
DB 13 CAATTTGTGGAACCTTCCGCTCAAGCCTTAGCGGTAAAGAACGCTGTGATTAAAGAGAC 72
QY 61 CAGCTCGATATCAGGAAGAAATGAAGCCAACTACAGGGAATGGCGAAGAGCTTTCT 120
DB 73 CAGCTCGATATCAGGAAGAAATGAAGCCAACTACAGGGAATGGCGAAGAGCTTTCT 132
QY 121 GAAATCATGCATGAGCA-----GATCTGCCCCCTGGAGAGAGACGAGCTTTAC 172
DB 133 GAAATCATGCATGAGCACTGCGGATGATGCCCCCTGGAGAGAGACGAGCTTTAC 192
QY 173 CGAATTCCTTCACATCTTCAACGCCATCACTGAGGACTCCACAAGACCAATGTTCTACG 232
DB 193 CGAATTCCTTCACATCTTCAACGCCATCACTGAGGACTCCACAAGACCAATGTTCTACG 252
QY 233 GGATGACCAAGCTGCTTTCGCTGATGATTACATCTCATGCCCCGTGTGGGACTTTC 292
DB 253 GGATGACCAAGCTGCTTTCGCTGATGATTACATCTCATGCCCCGTGTGGGACTTTC 312
QY 293 CTTTGCATTTGCCAATCCAGATGCTTTCCAAAGCCAACTACCTGGGAGACGAGCACA 352
DB 313 CTTTGCATTTGCCAATCCAGATGCTTTCCAAAGCCAACTACCTGGGAGACGAGCACA 372
QY 353 GGGAGACCAAGGGGAGAGGAGAGAAATGAAGCAACGTTATTCTTAAACAGA 412
DB 373 GGGAGACCAAGGGGAGAGGAGAGAAATGAAGCAACGTTATTCTTAAACAGA 432
QY 413 CTTTCTATAGAGATTGTAAGAGGTGCACATATTTTAAATCTCACTGGCAATATTC 472
DB 433 CTTTCTATAGAGATTGTAAGAGGTGCACATATTTTAAATCTCACTGGCAATATTC 492
QY 473 AAGTTTCATTTGCTCTTAAACAAAGTGT 501
DB 493 AAGTTTCATTTGCTCTTAAACAAAGTGT 521

RESULT 3
B1602169 777 bp mRNA linear EST 07-SEP-2001

LOCUS B1602169
DEFINITION B1602169 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5288423 5',
mRNA sequence.

ACCESSION B1602169
VERSION B1602169.1 GI:15495108
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 777)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgrahs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11729 row: d column: 24
High quality sequence stop: 777.
Location/Qualifiers
1..777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5288423"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 223 a 159 c 207 g 188 t
ORIGIN
Query Match 96.4%; Score 483; DB 13; Length 777;
Best Local Similarity 98.4%; Pred. No. 1.2e-115;
Matches 501; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 CAATTTGTGGAAGCTTCGGCTCAAGCCTTAGCGTAAACGACGCTGATTAAGAAGAC 60
|||||
DB 208 CAATTTGTGGAAGCTTCGGCTCAAGCCTTAGCGTAAACGACGCTGATTAAGAAGAC 267
QY 61 CAGCTCAGATATCAGAGAAATGAAGCCACTACAGGGAATGGCGAAGAGCTTTCT 120
|||||
DB 268 CAGCTCAGATATCAGAGAAATGAAGCCACTACAGGGAATGGCGAAGAGCTTTCT 327
QY 121 GAATTCATGATGAGCA-----GATCTGCCCTCGAGGAGGAAGAGCGGCTTTAC 172
|||||
DB 328 GAATTCATGATGAGCAAGCTGGGATGATCTGCCCTCGAGGAGGAAGAGCGGCTTTAC 387
QY 173 CGAATTCCTTCACATCTTCAAGCCATCAAGTGGACTCCCAACAGACAAATGTTTACG 232
|||||
DB 388 CGAATTCCTTCACATCTTCAAGCCATCAAGTGGACTCCCAACAGACAAATGTTTACG 447
QY 233 GGATGACACAGCTGCTTTCGGTGTGATTAATCTCATGCGCCGCTGTGGGACTTG 292
|||||
DB 448 GGATGACACAGCTGCTTTCGGTGTGATTAATCTCATGCGCCGCTGTGGGACTTG 507
QY 293 CTTTGTCAATTTGCAAACTCAGATGCTTTCCAAAGCCCAATCAGTGGGAGACGAGACA 352
|||||
DB 508 CTTTGTCAATTTGCAAACTCAGATGCTTTCCAAAGCCCAATCAGTGGGAGACGAGACA 567
QY 353 GGGAGAGCAAGGGGAGGAGAGAAAGAAATGAAGCAACAGTATTCTTAACAGA 412
|||||
DB 568 GGGAGAGCAAGGGGAGGAGAGAAAGAAATGAAGCAACAGTATTCTTAACAGA 627
QY 413 CTTTCTATAGAGATTGAAGAGTGCACATATTTTAAATCTCAGTGGCATATTC 472
|||||
DB 628 CTTTCTATAGAGATTGAAGAGTGCACATATTTTAAATCTCAGTGGCATATTC 687
QY 473 AAGTTTCATGTTGTCTTAACAAAGTGT 501

```

```

DB 688 AAGTTTCATGTTGTCTTAACAAAGTGT 716
|||||
RESULT 4
LOCUS AM162535/c
DEFINITION au77b01.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
sequence.
ACCESSION AM162535
VERSION AM162535.1 GI:6301568
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
Kizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Maira,M., Martin
White,Y., Wylie,T., Waterston,R. and Willson,R.
WashU-NCI human EST Project
Other ESTs: au77b01.y1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Possible reversed clone: PolyT not found
Seq primer: -400p from Gldco
High quality sequence stop: 455.
Location/Qualifiers
1..644
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2782249"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGCTCAAGATCTTAATTAATTAATCCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGCTCAAGATCTTAATTAATTAATTAATCCCCCCCCCCC-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."
BASE COUNT 162 a 168 c 128 g 186 t
ORIGIN
Query Match 95.8%; Score 479.8; DB 10; Length 644;
Best Local Similarity 98.0%; Pred. No. 8.3e-115;
Matches 499; Conservative 0; Mismatches 2; Indels 8; Gaps 1;
QY 1 CAATTTGTGGAAGCTTCGGCTCAAGCCTTAGCGTAAACGACGCTGATTAAGAAGAC 60
|||||
DB 638 CAATTTGTGGAAGCTTCGGCTCAAGCCTTAGCGTAAACGACGCTGATTAAGAAGAC 579
QY 61 CAGCTCAGATATCAGAGAAATGAAGCCACTACAGGGAATGGCGAAGAGCTTTCT 120
|||||

```

Db 578 CAGCTGAGTATGAGAGAAATGATAGCCAACTACAGGGAAATGGCGAAGACCTTTCT 519
Qy 121 GAATCATGATGAGCA-----GATCTGCCCCCTGGAGAGAGAGAGCGCTTAC 172
Db 518 GAATCATGATGAGCAGCGTGGATGATCTGCCCCCTGGAGAGAGAGAGCGCTTAC 459
Qy 173 CGAATTCCTTCATCTTCAACGCCATCAGTGGGACTCCAAACAGCAATGGTTCAG 232
Db 458 CGAATTCCTTCATCTTCAACGCCATCAGTGGGACTCCAAACAGCAATGGTTCAG 399
Qy 233 GGAATGACACGCTGCTGGGTGGTGAATTCATCTCAATGGCCCTGCTGGGACTTG 292
Db 398 GGAATGACACGCTGCTGGGTGGTGAATTCATCTCAATGGCCCTGCTGGGACTTG 339
Qy 293 CTTTGTATTTGCAAACTCAGATGCTTCCAAAGCCATCACTGGGAGACCGAGAC 352
Db 338 CTTTGTATTTGCAAACTCAGATGCTTCCAAAGCCATCACTGGGAGACCGAGAC 279
Qy 353 GGGAGAGCCAAAGGGGAGAGAGAGAAAGAAATAAGAACACGTTATTCTTAACAG 412
Db 278 GGGAGAGCCAAAGGGGAGAGAGAGAAAGAAATAAGAACACGTTATTCTTAACAG 219
Qy 413 CTTTGTATTTGAGAGTGTAAAGAGTGCACATATTTTAAATCTCAGCAATTTCA 472
Db 218 CTTTGTATTTGAGAGTGTAAAGAGTGCACATATTTTAAATCTCAGCAATTTCA 159
Qy 473 AAGTTTCAATTTGTCTTAACAAAGGTGT 501
Db 158 AAGTTTCAATTTGTCTTAACAAAGGTGT 130

RESULT 5
Bg676871 920 bp mRNA linear EST 01-MAY-2001
LOCUS 602623473F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748442 5',
DEFINITION mRNA sequence.
ACCESSION Bg676871
VERSION Bg676871.1 GI:13908268
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10600 row: a column: 19
High quality sequence stop: 849.
Location/Qualifiers
1..920

FEATURES
source
1..920

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4748442"
/clone_id="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NCI;
Site_2: Salti; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI/CGAP Library."
BASE COUNT 274 a 164 c 230 g 251 t 1 others
ORIGIN

Query Match 95.6%; Score 479; DB 12; Length 920;
Best Local Similarity 98.6%; Pred. No. 1.4e-114;
Matches 494; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
Qy 1 CAATTTGTGGAAGCTTGGGTCAGCCCTTACCGGTAAACGACCTGTGATTAAGAAGAC 60
Db 19 CAATTTGTGGAAGCTTGGGTCAGCCCTTACCGGTAAACGACCTGTGATTAAGAAGAC 76
Qy 61 CAGCTCGAGTGTGAGGAAGAAATGAAGCCATCAGAGGAAATGGCGAAGAGCTTTCT 120
Db 77 CAGCTCGAGTGTGAGGAAGAAATGAAGCCATCAGAGGAAATGGCGAAGAGCTTTCT 136
Qy 121 GAATCATGATGAGCAGATGCTGCCCTGGAGAGAGAGAGAGAGAGAGCTTTACGAATTC 180
Db 137 GAATCATGATGAGCAGATGCTGCCCTGGAGAGAGAGAGAGAGAGAGCTTTACGAATTC 196
Qy 181 CTTACATCTTCAAGCCCATCAGTGGGACTCCAAACAGCAATGTTTCAGGAGTACC 240
Db 197 CTTACATCTTCAAGCCCATCAGTGGGACTCCAAACAGCAATGTTTCAGGAGTACC 256
Qy 241 AGCTGCTGCTGGGTGGTGAATTCATCTCAATGGCCCTGCTGGGACTTGTCTCA 300
Db 257 AGCTGCTGCTGGGTGGTGAATTCATCTCAATGGCCCTGCTGGGACTTGTCTCA 316
Qy 301 TTTGCAAACTCAGATGCTTCCAAAGCCATCACTGGGAGACCGAGAGAGAGAC 360
Db 317 TTTGCAAACTCAGATGCTTCCAAAGCCATCACTGGGAGACCGAGAGAGAGAC 376
Qy 361 CAAGGGAGAGGGGAGAGAGAAATGAAGAACACGTTATTCTTAACAGACTTTCTAT 420
Db 377 CAAGGGAGAGGGGAGAGAGAAATGAAGAACACGTTATTCTTAACAGACTTTCTAT 436
Qy 421 AGGAGTGTAGAGAGTGCACATATTTTAAATCTCAGTGGCAATTTCAAGTTTC 480
Db 437 AGGAGTGTAGAGAGTGCACATATTTTAAATCTCAGTGGCAATTTCAAGTTTC 496
Qy 481 ATTGTCTTTAACAAGGTGT 501
Db 497 ATTGTCTTTAACAAGGTGT 517

RESULT 6
Bi603180 801 bp mRNA linear EST 07-SEP-2001
LOCUS 603249824F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301749 5',
DEFINITION mRNA sequence.
ACCESSION Bi603180
VERSION Bi603180.1 GI:15496119
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 801)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shifaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11763 row: p column: 06
High quality sequence stop: 710.
Location/Qualifiers
1..801

FEATURES
source
1..801
/organism="Homo sapiens"
/db_xref="taxon:9606"


```

/clone="IMAGE:5301749"
/clone.lib="NIH_MGC_96"
/tissue.type="Hypothalamus"
/lab.host="DH10B"
/Note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3';
size-selected for average insert size 2.3 kb and
normalized to R0r 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      233 a      164 c      215 g      189 t
ORIGIN

```

```

Query Match      87.3%; Score 437.2; DB 13; Length 801;
Best Local Similarity 96.3%; Pred. No. 1.1e-103;
Matches 493; Conservative 0; Mismatches 8; Indels 11; Gaps 4;

```

```

QY 1 CAATTGTGGAAGCTTGGGTCAAGCTTACGGTTAAACGAACGCTGATTAAAGAACAC 60
  |||||||
DB 208 CAATTGTGGAAGCTTGGGTCAAGCTTACGGTTAAACGAACGCTGATTAAAGAACAC 267
  |||||||
QY 61 CACCTGAGATCAGAGAAATGAAGCCACTACAGGAAATGGCGAGAGAGCTTCT 120
  |||||||
DB 268 CACCTGAGATCAGAGAAATGAAGCCACTACAGGAAATGGCGAGAGAGCTTCT 327
  |||||||
QY 121 GAATCATGATGAGCA-----GATCTGCCCTGGAGAGAGAGAGAGAGCTTTAC 172
  |||||||
DB 328 GAATCATGATGAGCAAGCTGGAGATGCTGCCCTGGAGAGAGAGAGAGAGCTTTAC 387
  |||||||
QY 173 CGAATTCCTTCACATCTTCAAGCCATCAGTGGGAGCTCCAAACAGCAATGGTTCACG 232
  |||||||
DB 388 CGAATTCCTTCACATCTTCAAGCCATCAGTGGGAGCTCCAAACAGCAATGGTTCACG 447
  |||||||
QY 233 GGATGACCAAGCTGCTTGGTGGTGTGATTCATCTCATGCGCCGCTGTGGGGAC-TT 291
  |||||||
DB 448 GGATGACCAAGCTGCTTGGTGGTGTGATTCATCTCATGCGCCGCTGTGGGGAC-TT 507
  |||||||
QY 292 GCTTGTCAATGGCAACCTCAGATGCTTCCAAAGCCATCAGTGGGAGAGAGAGCTT 351
  |||||||
DB 508 GCTTGTCAATGGCAACCTCAGATGCTTCCAAAGCCATCAGTGGGAGAGAGAGCTT 567
  |||||||
QY 352 AGGAGAGACCAAGGGAGAGAGAGAGAAATGAAGCAACAGCTATTCTTAAACAG 411
  |||||||
DB 568 AGGAGAGACCAAGGGAGAGAGAGAGAAATGAAGCAACAGCTATTCTTAAACAG 627
  |||||||
QY 412 ACTTGTATAGAGCTGTGAAGAGGTGCA-CAATTTTCTTAAATCTGA-CTGGCAATAT 469
  |||||||
DB 628 ACTTGTATAGAGCTGTGAAGAGGTGCA-CAATTTTCTTAAATCTGA-CTGGCAATAT 687
  |||||||
QY 470 TCAAGATTTCATTTGTCTTAAACAAGGTGT 501
  |||||||
DB 688 TCAAGATTTCATTTGTCTTAAACAAGGTGT 719
  |||||||

```

```

RESULT 7
LOCUS      BF748808      546 bp      mRNA      linear      EST 10-JAN-2001
DEFINITION MRO-BN0115-041000-013-e10 BN0115 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF748808
VERSION    BF748808.1 GI:12075471
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens

```

```

REFERENCE 1 (bases 1 to 546)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva W. Jr., Zago,M.A., Bordini,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
          Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

```

```

TITLE      'M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
JOURNAL    Contact: Simpson A.J.G.
MEDLINE    Laboratory of Cancer Genetics
COMMENT     Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?PL=MRO&LZ=MRO-BN0115-
            041000-013-e10&t3=2000-10-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 16
            High quality sequence stop: 545.

```

FEATURES

```

source     .
            location/Qualifiers
            1..546

```

```

/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone.lib="BN0115"
/dev.stage="Adult"

```

```

/Note="Organ: breast,normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

BASE COUNT      140 a      135 c      115 g      156 t
ORIGIN

```

```

Query Match      85.3%; Score 427.2; DB 12; Length 546;
Best Local Similarity 99.1%; Pred. No. 4.5e-101;
Matches 440; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

QY 59 ACCAGCTCCAGTATCAGAGAAATGAAGCAACTACAGGAAATGGCGAGAGAGCTT 118
  |||||||
DB 546 ACCAGCTCCAGTATCAGAGAAATGAAGCAACTACAGGAAATGGCGAGAGAGCTT 487
  |||||||
QY 119 CTGAATTCATGCTAGAGCATCTGCCCTGGAGAGAGAGAGAGAGCTTACGAAAT 178
  |||||||
DB 486 CTGAATTCATGCTAGAGCATCTGCCCTGGAGAGAGAGAGAGAGCTTACGAAAT 427
  |||||||
QY 179 CCTTTCACATCTTCAAGCCATCAGTGGAGCTCCAAACAGCAATGTTTACGGATGA 238
  |||||||
DB 426 CCTTTCACATCTTCAAGCCATCAGTGGAGCTCCAAACAGCAATGTTTACGGATGA 367
  |||||||
QY 239 CCAAGCTCGTTCGCGTGTGATTCATCTGAGCCCGTGTGGAGAGCTTGTGT 298
  |||||||
DB 366 CCAAGCTCGTTCGCGTGTGATTCATCTGAGCCCGTGTGGAGAGCTTGTGTGT 307
  |||||||
QY 299 CATTTGCAAACTCAGAGATCTTCCAAAGCAATCAGTGGGAGAGAGAGAGAGG 358
  |||||||
DB 306 CATTTGCAAACTCAGAGATCTTCCAAAGCAATCAGTGGGAGAGAGAGAGAGG 247
  |||||||
QY 359 ACCAAGGGAGAGGAGAGAAATGAAGCAACAGTAAATTT-CTTAACAGACTTTC 417
  |||||||
DB 246 ACCAAGGGAGAGGAGAGAAATGAAGCAACAGTAAATTT-CTTAACAGACTTTC 187
  |||||||
QY 418 TATAGAGTTGTGAAGGTGACATATTTTAAATCTCAGTGGCAATATTCAAACTT 477
  |||||||
DB 186 TATAGAGTTGTGAAGGTGACATATTTTAAATCTCAGTGGCAATATTCAAACTT 127
  |||||||
QY 478 TTCAATGTGTCTTAAACAAGGTGT 501
  |||||||
DB 126 TTCAATGTGTCTTAAACAAGGTGT 103
  |||||||

```

```
RESULT 8
AM953240          679 bp      mRNA      linear      EST 01-JUN-2000
LOCUS             AM953240
DEFINITION        ES3365310 MAGE resequences, MAGB Homo sapiens cDNA, mRNA sequence.
ACCESSION         AM953240
VERSION           AM953240.1  GI:8142923
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 679)
AUTHORS           Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
                  ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
                  Quackenbush,J.
TITLE             Assessment of gene expression patterns in a model of colon tumor
                  metastasis using a 19,200 element cDNA microarray
JOURNAL           Unpublished (2000)
COMMENT           Contact: John Quackenbush
                  The Institute for Genomic Research
                  9712 Medical Center Dr., Rockville, MD 20850, USA
                  Tel: 301 838 3528
                  Fax: 301 838 0208
                  Email: johnq@tigr.org
                  Plate: 48
                  Seq primer: Reverse.
FEATURES          source          1..679
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_lib="MAGE resequences, MAGB"
                  /note="Vector: pBluescriptSkm"
BASE COUNT       194 a      146 c      191 g      147 t      1 others
ORIGIN
Query Match      82.0%; Score 410.8; DB 10; Length 679;
Best Local Similarity 97.5%; Pred. No. 8.8e-97;
Matches 430; Conservative 0; Mismatches 3; Indels 8; Gaps 1;
QY 1 CAAATTGTGGAAGCTTGGCGTCAAGCCTTAGCGGTAACGAACGCTGTGATTAAGACAGAC 60
DB 239 CAATTGTGGAAGCTTGGCGTCAAGCCTTAGCGGTAACGAACGCTGTGATTAAGACAGAC 298
QY 61 CAGCTCGAGTATCAGGAGAAATGAAGCCAACTACAGGGAATGCGCAAGAGAGCTTCT 120
DB 299 CAGCTCGAGTATCAGGAGAAATGAAGCCAACTACAGGGAATGCGCAAGAGAGCTTCT 358
QY 121 GAATCATCATGAGCA-----GATCTGCCCTTGAGAGAGAAGACGATCTTAC 172
DB 359 GAATCATCATGAGCAGCTGGAGTATCTGCCCTTGAGAGAGAAGACGATCTTAC 418
QY 173 CGAATTCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAGCACATGTTACG 232
DB 419 CGAATTCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAGCACATGTTACG 478
QY 233 GGAATGACAGCTGCTGGTGGTGTGATTAATCTCATGCGCGGTGTGGGACTTG 292
DB 479 GGAATGACAGCTGCTGGTGGTGTGATTAATCTCATGCGCGGTGTGGGACTTG 538
QY 293 CTTTGATATTTGCAAACTCAGATGCTTCCAAAGCCAATCACTGGGAGACGAGACA 352
DB 539 CTTTGATATTTGCAAACTCAGATGCTTCCAAAGCCAATCACTGGGAGACGAGACA 598
QY 353 GGGAGACCAAGGGGAGAGAGAAATTAAGACAACGTTATTTCTTAACAGA 412
DB 599 GGGAGACCAAGGGGAGAGAGAAATTAAGACAACGTTATTTCTTAACAGA 658
QY 413 CTTTCTATAGAGTGTGAAGA 433
DB 659 CTTTCTTATAGAGTGTGAAGA 679
```

```
RESULT 9
BF667219          574 bp      mRNA      linear      EST 21-DEC-2000
LOCUS             BF667219
DEFINITION        602121018r1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278077 5',
                  mRNA sequence.
ACCESSION         BF667219
VERSION           BF667219.1  GI:11941114
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 574)
AUTHORS           NIH-MGC http://mgc.nci.nih.gov/
                  National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE             Unpublished (1999)
JOURNAL           Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
COMMENT           CDNA library Preparation: ATCC
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: CLONTECH Laboratories, Inc.
                  CDNA Sequencing by: Incyte Genomics, Inc.
                  DNA distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLCM1101 row: c column: 06
                  High quality sequence stop: 544.
FEATURES          source          1..574
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_lib="IMAGE:4278077"
                  /clone_lib="NIH_MGC_56"
                  /tissue_type="orimlative neuroectoderm"
                  /lab_host="DH10B (T1 phage-resistant)"
                  /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                  SfiI (ggcgccgtcgcc); Site_2: SfiI (ggccattatggcc);
                  Double-stranded cDNA was prepared from cell line RNA. 5'
                  and 3' adaptors were used in cloning as follows: 5'
                  adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGCCGACATG-3' (30)BN-3'
                  sequence: 5'-ATTCTAGAGCGGAGCGCGCCGACATG-3' (30)BN-3'
                  (where B = A, C, G, or T). Average
                  insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
                  contained inserts by PCR. This library was enriched for
                  full-length clones and was constructed by Clontech
                  Laboratories (Palo Alto, CA)."
BASE COUNT       165 a      133 c      158 g      118 t
ORIGIN
Query Match      76.9%; Score 385.2; DB 12; Length 574;
Best Local Similarity 99.2%; Pred. No. 4.2e-90;
Matches 387; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 ATTTGTGGAAGCTTGGCGTCAAGCCTTAGCGGTAACGAACGCTGTGATTAAGACAGCA 62
DB 185 AATTTGTGGAAGCTTGGCGTCAAGCCTTAGCGGTAACGAACGCTGTGATTAAGACAGCA 244
QY 63 GCTCGAGTATCAGGAGAAATGAAGCCAACTACAGGGAATGCGCAAGAGAGCTTCTGA 122
DB 245 GCTCGAGTATCAGGAGAAATGAAGCCAACTACAGGGAATGCGCAAGAGAGCTTCTGA 304
QY 123 AATCATGATGAGCAGATCTGCCCTTGAGAGAGACGAGCTTACCGAATTCCT 182
DB 305 AATCATGATGAGCAGATCTGCCCTTGAGAGAGACGAGCTTACCGAATTCCT 364
QY 183 TCACATCTTCAAGCCCATCAGTGGAGATCCCAACAAGACACATGTTTACGGGATGAGCAG 242
DB 365 TCACATCTTCAAGCCCATCAGTGGAGATCCCAACAAGACACATGTTTACGGGATGAGCAG 424
QY 243 CTCGCTTCGGTGGTGGATTAACATCATATGAGCCCGTGTGGGAGACTTCTTGTCAAT 302
DB 425 CTCGCTTCGGTGGTGGATTAACATCATATGAGCCCGTGTGTGGGAGACTTCTTGTCAAT 484
```

QY	303	TGC	AACTCAGATGATCTCTTCCAAAGCCCAATCCTGGGAGACCGAGCACAGGAGAGACCA	362
QY	485	TGC <td>AACTCAGATGATCTCTTCCAAAGCCCAATCCTGGGAGACCGAGCACAGGAGAGACCA</td> <td>544</td>	AACTCAGATGATCTCTTCCAAAGCCCAATCCTGGGAGACCGAGCACAGGAGAGACCA	544
Db	363	AGGGGAAGGGAGAGCAAAAGGAATATAAGAA	392	
QY	545	AGGGGAAGGGAGAGCAAGGAATATACGAA	574	

RESULT 10	BM353207	521 bp	linear	EST 07-JAN-2002
LOCUS	BM353207			
DEFINITION	I644a10.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:063603			
ACCESSION	BM353207			
VERSION	BM353207.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE AUTHORS

TITLE	Endocrine Pancreas Consortium
JOURNAL	Unpublished (2000)
COMMENT	Other_ESTS: 1g44a10.x1

Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmclinton@molp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from G1pco
High quality sequence stop: 462.

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH08"
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alon Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@metgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

```

BASE COUN
ORIGIN

Query Match	73.78;	Score 369;	DB 13;	Length 521;
Best Local Similarity	100.0%;	Pred. No. 7.1e-86;		
Matches 369;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 1 CAATTGTTGGAAAGCTTGGCGGTCAAGCCTTAGCGGTAAACGAACGTCGATTTAAGAGAGAC 60
 |||||
 153 CAATTGTTGGAAAGCTTGGCGGTCAAGCCTTAGCGGTAAACGAACGTCGATTTAAGAGAGAC 212
 Oy 61 CACCTGCAGCTATTCAGAGAGAAATGAAAGCCACACTACAGGAGAAATGGCCAGAGAGCTTTCT 122

Db	213	CACGTGAGATATCAGGAAGAAATGAAACCCACACTACAGGGAAATGGCGAAGGAGCTTCT	272
Qy	121	GAATCATGATGATGAGAGATCTGCCCCCTGGAGAGAAAGAGCGCTTTACCGAATTC	180
Db	273	GAATCATGATGATGAGAGATCTGCCCCCTGGAGAGAAAGAGCGCTTTACCGAATTC	332
Qy	181	CTTCACATCTTCACAGCCCATCATGATGGAGCTCCAAACAAGCAATGTGTTACGGATGACC	240
Db	333	CTTCACATCTTCACAGCCCATCATGATGGAGCTCCAAACAAGCAATGTGTTACGGATGACC	392
Qy	241	AGCTCGCTTCGGCTCGTGTGATTACATCTTCATGGCCCCGCTGTGTGGGAGACTTGCCTTTTCA	300
Db	393	AGCTCGCTTCGGCTCGTGTGATTACATCTTCATGGCCCCGCTGTGTGGGAGACTTGCCTTTTCA	452
Qy	301	TTTGCAAACTCAGGATGCTTTTCCAAGGCATTCATCTGGGAGACCGGACACAGGAGAGAC	360
Db	453	TTTGCAAACTCAGGATGCTTTTCCAAGGCATTCATCTGGGAGACCGGACACAGGAGAGAC	512
Qy	361	CAAGGGAA 369	
Db	513	CAAGGGAA 521	

RESULT 11	
BM353362	
LOCUS	539 bp mRNA linear EST 07-JAN-2002
DEFINITION	i946410.v1 HR85 islet Homo sapiens CDNA 5' similar to TR:O65603
ACCESSION	O65603.FRG PROTEIN ; , mRNA sequence.
VERSION	BM353362
KEYWORDS	BM353362.1 GI:18085720
SOURCE	EST.
	human.

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 539)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

TITLE	Endocrine Pancreas Consortium
JOURNAL	unpublished (2000)
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Email: dmelton@iobp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@wustl.edu)
Seq primer: -40RP from Glbco
High quality sequence stop: 478.

FEATURES

SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85_islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/notice="Organ: Pancreas; Vector: pBluescript SK(-); Site:1:
NotI; Site:2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism DIV, (Alan Fernutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.

```

Louis, MO 63110, E-mail: hinoe@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692.
BASE COUNT 153 a 127 c 144 g 114 t 1 others

Query Match 73.5%; Score 368; DB 13; Length 539;
Best Local Similarity 99.7%; Pred. No. 1,3e-85;
Matches 368; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAATTGTGGAGCTTGGCGTCAAGCCTTACGGGTAAGCAGCTGTATTAAAGAGAC 60
|||||
Db 171 CAATTGTGGAGCTTGGCGTCAAGCCTTACGGGTAAGCAGCTGTATTAAAGAGAG 230
CAGCTCGACTATCAGAGAAATGAAGCCACTACAGGAAATGGGAGAGAGCTTTCT 120
|||||
Db 231 CAGCTCGACTATCAGAGAAATGAAGCCACTACAGGAAATGGGAGAGAGCTTTCT 290
QY 121 GAATTCATGATGACAGACATCTGCCCTGGAGAGAGACAGCTCTTACCGAATTCC 180
|||||
Db 291 GAATTCATGATGACAGACATCTGCCCTGGAGAGAGACAGCTCTTACCGAATTCC 350
CTTCACATCTTCAACGCCATCAGTGGAGCTCCACAAAGCAACATGGTTACGGGATGAC 240
|||||
Db 351 CTTCACATCTTCAACGCCATCAGTGGAGCTCCACAAAGCAACATGGTTACGGGATGAC 410
QY 241 AGCTGCTCTTGGTGTGTGATTCATCTCATGCCCCGTGTGTGGGACTTGTCTTGTCA 300
|||||
Db 411 AGCTGCTCTTGGTGTGTGATTCATCTCATGCCCCGTGTGTGGGACTTGTCTTGTCA 470
QY 301 TTTCGAACTCGAGATGCTTTCCAAAGCCATCACTGGGGAGACCGAGCAGAGAGAGAC 360
|||||
Db 471 TTTCGAACTCGAGATGCTTTCCAAAGCCATCACTGGGGAGACCGAGCAGAGAGAGAC 530
QY 361 CAAGGGGAA 369
|||||
Db 531 CAAGGGGAA 539

RESULT 12 359 bp mRNA linear EST 15-AUG-1997
AA484945
LOCUS aa41b06.r1 NCI-CGAP-GCB1 Homo sapiens CDNA clone IMAGE:815795 5
DEFINITION similar to TR:G550420 G550420 TRG mRNA ; mRNA sequence.
ACCESSION AA484945
VERSION AA484945.1 GI:2214164
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 359)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgep@strahl.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/HLN at:
www.bio.lnll.gov/bdip/image/image.html
Seq primer: -28ml3 rev1 ET from AmerSham.

FEATURES

source
1..359
/organism="Homo sapiens"
/db_xref="GDB:6035413"
/db_xref="taxon:9606"
/clone="IMAGE:815795"

/clone.lib="NCI-CGAP-GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - oligo(dt) primer
5'-TGTACCAATCTGAAGGAGGAGGCGGCGCTCATTTTCTTTTCTTTT-3'
1. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 108 a 80 c 97 g 74 t
ORIGIN
Query Match 71.7%; Score 359; DB 9; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.8e-83;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 TGAATTAAAGAGACACGCTCGACTATCAAGAAATGAAGCCAACTACAGGGAAATGG 106
|||||
Db 1 TGAATTAAAGAGACACGCTCGACTATCAAGAAATGAAGCCAACTACAGGGAAATGG 60
QY 107 CGAAGAGCTTTCTGAAATCATGATGAGCAGATCTGCCCTGGAGAGAGACGAGCG 166
|||||
Db 61 CGAAGAGCTTTCTGAAATCATGATGAGCAGATCTGCCCTGGAGAGAGACGAGCG 120
QY 167 TCTTACCGAATTCCTTCACATCTTCAACGCCATCAGTGGAGCTCCAAAGCAACATGG 226
|||||
Db 121 TCTTACCGAATTCCTTCACATCTTCAACGCCATCAGTGGAGCTCCAAAGCAACATGG 180
QY 227 TTCACGGGATGACACAGCTCGTTCGTGTGTGTGATTACATCTCATGCCCCGTGTGTGG 286
|||||
Db 181 TTCACGGGATGACACAGCTCGTTCGTGTGTGTGATTACATCTCATGCCCCGTGTGTGG 240
QY 287 GACTTGCTTTGTGATTTGGAAACTCAGATGCTTCCAAAGCCAACTACAGGGAGACCG 346
|||||
Db 241 GACTTGCTTTGTGATTTGGAAACTCAGATGCTTCCAAAGCCAACTACAGGGAGACCG 300
QY 347 AGCACAGGAGAGACCAAGGGAGGAGAGAAAGAAATGAAGACACCTTATTTCT 405
|||||
Db 301 AGCACAGGAGAGACCAAGGGAGGAGAGAAAGAAATGAAGACACCTTATTTCT 359

RESULT 13 609 bp mRNA linear EST 13-JUN-2001
BG992490
LOCUS MR3-HT1039-020201-002-b09 HT1039 Homo sapiens CDNA, mRNA sequence.
DEFINITION BG992490
ACCESSION BG992490.1 GI:14396560
VERSION BG992490.1 GI:14396560
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 609)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordh,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsumura,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

TITLE
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:39:15 ; Search time 994.674 Seconds

(without alignments)
14629.322 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500

Perfect score: 500

Sequence: 1 agtttacacccatcaccaaa.....agggcagctgtagtcgcct 500

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
GenBank:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_hum:*
32: em_htg_hum:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498.4	99.7	7506	6	AX255048
2	498.4	99.7	7522	9	AF527605
3	498.4	99.7	7545	9	AB028981
4	204.4	40.9	6454	6	AX173022
5	194.8	39.0	139887	9	CNS01RGX
6	126.8	25.4	163316	9	AL161420
7	90.4	18.1	175281	2	AC109966
8	85	17.0	4391	6	AX173028
9	85	17.0	4393	6	AX173118
10	75.8	15.2	175281	2	AC109966
11	75.8	15.2	192825	2	AC119357
12	71.4	14.3	5564	9	AK074081
13	71.4	14.3	6422	9	AK090429
14	71.4	14.3	7215	6	AX174569
15	71.4	14.3	9298	6	HSMB03577
16	69.8	14.0	665	6	AX371133
17	60.8	12.2	181127	10	AL672038
18	60.2	12.0	58723	2	AL391237
19	59	11.8	162378	2	AC105541
20	59	11.8	174007	2	AC118998
21	58.2	11.6	82603	2	AC111850
22	52.8	10.6	171811	9	AC011739
23	52	10.4	181179	9	AL161725
24	48.4	9.7	167215	5	AY016023
25	48	9.6	281	11	G47605
26	46	9.2	4886	9	AB037816
27	46	9.2	6372	6	AX173175
28	43.4	8.7	82603	2	AC111850
29	42.6	8.5	147556	2	AC110359
30	42.4	8.5	73515	2	AC011735
31	42.4	8.5	166626	3	AC008318
32	42.4	8.5	261475	2	AC096357
33	42.4	8.5	301639	3	AE003590
34	41	8.2	6828	6	AX172874
35	40.4	8.1	211	6	AX173251
36	40.4	8.1	3023	9	AK022412
37	40.4	8.1	122321	9	AC011472
38	38.6	7.7	69694	2	AC020073
39	38.6	7.7	156222	2	AC087224
40	38.6	7.7	159108	2	AC026031
41	38.6	7.7	159350	3	AC008354
42	38.6	7.7	257224	3	AE003617
43	37.8	7.6	181994	2	AC116703
44	37.6	7.5	154846	2	AP001374
45	37.6	7.5	156211	9	AC091551

ALIGNMENTS

RESULT 1	AX255048	7506 bp	DNA	Linear	PAT 10-OCT-2001
LOCUS	AX255048				
DEFINITION	Sequence 7 from Patent WO0170808.				
ACCESSION	AX255048				
VERSION	AX255048.1	GI:16074541			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	1 (bases 1 to 7506)				
TITLE	Rastelli, L.K. and Gerlitsen, M.				
	Angiogenesis-associated proteins, and nucleic acids encoding the				
	same				

JOURNAL	Patent: WO 0170808-A 7 27-SEP-2001; Curegen Corporation (US); GENENTECH, INC. (US)
FEATURES	source 1. .7506 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	2147 a 1689 c 1764 g 1906 t
ORIGIN	
Query Match	99.7%; Score 498.4; DB 6; Length 7506;
Best Local Similarity	99.8%; Pred. No. 3.7e-146;
Matches 499; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 AGTTTACACCATCACCACCAACCCAGAAATTTATGATGAGATTAAATAGATTGCCAC 60
DB	2142 AGTTTACACCATCACCACCAACCCAGAAATTTATGATGAGATTAAATAGATTGCCAC 2201
QY	61 TCAGTCGATGAAAAGCACCACCTGTTGTGCATTTCTTCATGTGACGTGACAACTC 120
DB	2202 TCAGTCGATGAAAAGCACCACCTGTTGTGCATTTCTTCATGTGACGTGACAACTC 2261
QY	121 AAGTAAAGAGACACGAGAGAGGATGTCGTTGAAACCCAAAGTGGCTACTCGTGC 180
DB	2262 AAGTAAAGAGACACGAGAGAGGATGTCGTTGAAACCCAAAGTGGCTACTCGTGC 2321
QY	181 TCCCTCTCTGAAAGACGGAAGGTGTGCAACAGCAGCAGCATCCGGTCTGGCGAA 240
DB	2322 TCCCTCTCTGAAAGACGGAAGGTGTGCAACAGCAGCAGCATCCGGTCTGGCGAA 2381
QY	241 CCTTCCTTGGGGCTATCTGTTGCTACCAAGAGCTTGGGATGGGACAGGCAATTATGTC 300
DB	2382 CCTTCCTTGGGGCTATCTGTTGCTACCAAGAGCTTGGGATGGGACAGGCAATTATGTC 2441
QY	301 AATTAAATGGGTAGATGAGAGGCAACCACTGCTGAAAAATTCACATCTGATTTCTAC 360
DB	2442 AATTAAATGGGTAGATGAGAGGCAACCACTGCTGAAAAATTCACATCTGATTTCTAC 2501
QY	361 AATGTAATCTCAGGATCAGCATTTACATTAATTTTTTCCAGTACTGTGCAAAACCGAATC 420
DB	2502 AATGTAATCTCAGGATCAGCATTTACATTAATTTTTTCCAGTACTGTGCAAAACCGAATC 2561
QY	421 TGGAGCCCAAGCCTTAGGAAACGAACTGTAAAGTACTTAAGTCTGCATGGATGGA 480
DB	2562 TGGAGCCCAAGCCTTAGGAAACGAACTGTAAAGTACTTAAGTCTGCATGGATGGA 2621
QY	481 AGGCCAGCTGATGATCGCCT 500
DB	2622 AGGCCAGCTGATGATCGCCT 2641
RESULT 2	
AF527605	7522 bp mRNA Linear PRI 01-AUG-2002
LOCUS	AF527605
DEFINITION	Homo sapiens zizimin1 mRNA, complete cds.
ACCESSION	AF527605
VERSION	AF527605.1 GI:22038158
KEYWORDS	human.
SOURCE	
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Meller,N., Iranl-Tehrani,M., Kiosses,W.B., Del Pozo,M.A. and Schwartz,M.A.
TITLE	zizimin1, a novel Cdc42 activator, reveals new guanine nucleotide exchange-exchange factor domain for rho proteins
JOURNAL	Nat. Cell Biol. (2002) In press
REFERENCE	2 (bases 1 to 7522)
AUTHORS	Meller,N. and Schwartz,M.A.
TITLE	Direct Submission
JOURNAL	Submitted (05-JUL-2002) Cell Biology, The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES	Location/Qualifiers

```

source
1. 7522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
56.6265
/!note="guanine nucleotide exchange factor: Cdc42
activator"
/codon_start=1
/product="zizimin1"
/protein_id="AA090306.1"
/db_xref="gi:22038159"
/translat="MSOPPLPASAETRTKPTALSKDGTAAELRQSEVYRGSVLLA
KPLELPEYENICITYOKTQTLINDCLREMLSEFPDQKQTLIRROGTCSTYPAKE
EASGLFEVETIKYNSDMWLVNFKYKREYDSEFPLDNDKQVLLKRLPVHYKVEPEVD
KDEAASISGOKGGITKRWGLXYKGNNAVSATSVTRSKRRRPHILIQLDGSYNLNRK
DEKISPEPKGSIPLDSCMGVQNNKRVRAPELMDQSRSTLLAADSVEEMETITLL
NKLIQINFEAAMQERNDSHEDEQSLBESGGSDLYPELAKSAREELIKLSBS
RVKLEFLYDPAQKLDIFSAAEYKSFEEKRGRKILVVCNPLSVCNOCVANEPEGPT
NMEPEPVLPLDIKYNRKISADPHVBLNHSYQMLATSPALMNSGQSPYLKGI
LHEAAQYRKQGISVTCPHPIDIFLVAIEKVLQGSITTHCAEYPMKSDSKVAOKVL
KNAQACORGLQYRMPFAMARTLFKAQSGNLDKNARFSAIYRQDSNKLSNDMLKLL
ADPRPEKMAKLPVILGMDITTDNDSNDEPNYVNSYIPKQETSKRPTITPEVEE
FVPCRRQTPYTYTNHLIYYPRYKLYDKDSKPARKNRAICIERDSEDESDPLK
CTIGPBGVPTNSFAFVLLHHNHPFYEYIKLELPTQJHEHNLILTFPHVCSNS
SKGTSKRDVETOVGYSWLPLKQGRVYSEQIIPYSAMPSGYLOELGKMGHGR
PEIKRWGDKPLKISTHLYSTVYTOQDHLNHFQYQCKTESGQALGNELVLYKSL
HAMEGHVIAFLPTILNQLFVLTRAQEEVAVNVTIIVHVAQCHEELESRLSR
VKYAKAEPPYASIEYKTVHEELTKSMITILKPSADELTSNKLKYSNPFEDVLIKSA
OHLIENSKVKLLRNQRPASVHNHAEPTVNMFLMTHQKRPDRPEAKNHNHSLAFLI
KRCFTPDNRGVPFOINNYISCFAPGDPKLTPEKFEFLVQCHNETIPALNMPBG
KGRIOYDQDLDYSLTDEFCRNFELVGLLREVTALQJEFREVLRAISVLNLIK
HSFDPRYASRSHQARIATLYLPLEGLLLENVQRLNVNADSPFNAMGTIVKDESLAP
AVNPLVTPQKSTLDNSLHKMLGALIGIASPYTSPNINSVYNADSRSLISTDGS
NLSBERNSEKNSLDKHOOSTGTGNSVYRCDKLDQSEIKELMCPYILKMSDNLAF
TYMNAKASTSELMDEPFTISEVCIHQFOYGRKRYATLRQNMHARILQOGLDNLSTFPHS
YGHSDADVLHOSLLEANIATVECVCTALDITSLPTARKNOLLADHGHNPLMKVPEVY
LCFTQKNOSETALKNVTFTALBSLTYKPTPTBGRADMCALCEIQLKCNSSLSIR
TEASQLEFLMRNNEPDTGKSKSEVRHLOVITVSOLADAVOIGTRPOQSLSIINN
CANSDRILKHTSFSSDVKDLTKRLIRTYLMAKQKBEHNDPEMIVDQYSLAKSVAST
PELKRSLDMSMARLHVKNGLDSEAKMYVATLVALEYLRKRAVOMEPLPLPHSISA
CLRSRGVGFQGCCTAERVITPNIDEASAMEDQVAFNEDVWMLLELLOCCADGTA
KAEREYLADIADYKLIPIYERKRPDERLALVYTHIHMAYDVFENWMLLELLOCCADGTA
VAFPGAOYQFQDSETFDVEGFDEDEGQKRYIKKEPLRPLSEISQRLKLYSDKGS
ENVMIDDSGKVRKDKSKYATIQYTHVITPFDEKLORKTEFESHNIIRFEM
PFTQYGRQGGVEQCKRRTITVAIHCEPVRKRILPMQOHNDLNEIVALIDEMSK
VAELRQKSAEVDMLKLOLQGSVQVYNAQPLAYARAFELDDTNTKTRPDKVALL
KEYRQFVEVACQGLAVANERLIKEDQLEYQEEKMYANREMAKELSETIMEQJCS"
BASE COUNT      2146 a      1694 c      1774 g      1908 t
ORIGIN
Query Match      99.7%: Score 498.4; DB 9; Length 7522;
Best Local Similarity 99.8%: Pred. No. 3,7e-146;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 AGTTTACACATCCCAAAACCCAGATTTTATGATGAGATTAAATAGAGTGGCCAC 60
2155 AGTTTACACCATCCCAAAACCCAGATTTTATGATGAGATTAAATAGAGTGGCCAC 2214
61 TCAGTCGATGAAAAGCACCACTGTGTGTCACATCTCTCCATGTCAGCTGTGACAATC 120
2215 TCAGTCGATGAAAAGCACCACTGTGTGTCACATCTCTCCATGTCAGCTGTGACAATC 2274
121 AAGTAAAGGAAGCAGGAAGAGGAGATGCTGTGAANAACCAATGGCTACATCCCTGGCT 180
2275 AAGTAAAGGAAGCAGGAAGAGGAGATGCTGTGTGAANAACCAATGGCTACATCCCTGGCT 2334
181 TCCCCTCCTGAAAAGCAGGAAGGATGTGATCAAGCGACACACATCCCGCTTCGGCGAA 240
2335 TCCCCTCCTGAAAAGCAGGAAGGATGTGATCAAGCGACACACATCCCGCTTCGGCGAA 2394
241 CCTTCCTTCGGGGCTATCTTGGCTTACCAAGACCTTGGATGGGACGACATTATGCTCCGGA 300
2395 CCTTCCTTCGGGGCTATCTTGGCTTACCAAGACCTTGGATGGGACGACATTATGCTCCGGA 2454

```


QY 301 AATTAAATGGGTAGATGAGGCAAGCCACTGCTGAAATTTCCACATCTGTTTCTAC 360
|||||
Db 2455 AATTAAATGGGTAGATGAGGCAAGCCACTGCTGAAATTTCCACATCTGTTTCTAC 2514
QY 361 AGTGTATCTACAGATCAGCATTTACATTAATTTTTCAGTACTGTGAGAAACCGAATC 420
|||||
Db 2515 AGTGTATCTACAGATCAGCATTTACATTAATTTTTCAGTACTGTGAGAAACCGAATC 2574
QY 421 TGGAGCCCAAGCCTTGGAAACGAACCTTGAAGTACCTTAAGAGTCGATGATGGA 480
|||||
Db 2575 TGGAGCCCAAGCCTTGGAAACGAACCTTGAAGTACCTTAAGAGTCGATGATGGA 2634
QY 481 AGGCACGTCGATGATGCTCCT 500
|||||
Db 2635 AGGCACGTCGATGATGCTCCT 2654

RESULT 3
AB028981 7545 bp mRNA linear PRI 10-MAY-2002
LOCUS Homo sapiens mRNA for KIAA1058 protein, partial cds.
ACCESSION AB028981
VERSION AB028981.2 GI:20521745
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pbuescriptII SK plus
clone:hl12146s1.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kikuno, R., Nagase, T., Ishikawa, K., Hirose, M., Miyajima, N.,
Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.
TITLE Prediction of the coding sequences of unidentified human genes.
XIV. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 6 (3), 197-205 (1999)
MEDLINE 10470851
PUBMED 99397452
REFERENCE 2 (bases 1 to 7545)
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
TITLE Direct Submission
SUBMITTED (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
COMMENT On May 9, 2002 this sequence version replaced gi:5689452.
FEATURES
Source Location/Qualifiers
1..7545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="hl12146s1"
/tissue="brain"
/clone_lib="pbuescriptII SK plus"
/note="This sequence was obtained by subcloning of the DNA
fragments derived from two cDNA clones (1 - 1846 was
derived from ef00877 and 1847 - 7545 was derived from
hl12146)."
1..7545
/gene="KIAA1058"
<1..6288
/gene="KIAA1058"
/codon_start=1
/product="KIAA1058 protein"
/protein_id="BA83010.2"
/db_xref="GI:20521745"
/translation="LACICLAPGTRSCSSRLSDLRQPRMQADKCRTSRSYKKE
LVESPLQYDAAGVEAEASPGVPAPKPLILPDIYENYIVKTKQIINDICREMLIF
PYDDPOTAILRQGRYICSTVPAKAEESQLEFCEIKTYISDMHLVYKEDYSGE
FRLPKYVKDKLPHYHVEYDEVDKDDASISQSGGIRKHGMIVKGNNSAISV
TKMSPKRRFFHILQDGSYINLNFYKDEKISKEPFSITLDCMKCYOVNNKRYRPAE
LKMDDKSYLLAADSEVEEMWITTLKTLQINPRAAOKERKNGSDSHEDDEOSKLEGS

GSGLDSTYLPDLAKSAREAEIKLSESRYKLFYLDPAOKLDFESSAEPEVKSFEERK
RIIVKQNDLSFNLCQCAVNEBEGPTINVEFFETLSLPIKYNRKISAPFADLHNS
VRQMLATSPALNMGSSQSPSYKGLIHNEAAQYPROGJLEFPCPLDFLVARLHVS
LOGSLTRCAIPYMKSSDSSKVAOKYKKNKQACORLGOVBMPEAMARLTFDASGNL
DKARSRAITRQDSNKLSDNDMLKLADRRKPEKMAKPLVITLGNLDTIDNYSDFPN
YVNSSTYPTKQFETCSKSTPTFEVEEFVPTKPHQTPYITNHLVYKTYLKYDSK
SFAKARNIAICIEFKDSEDSQPLCIYGRPGPVPYTSFAVAALHHQNDPEYDEI
KIELPDQLHRKHHLLTFEPVSCDSSKSGTGRDVEYFOVGSMLPLDKDRVYSE
OHIPVANSPLPSGYLGOELGMRHYCEPKLWVNGKPLIKLSTHVSITYTODQHLN
FPOYCSQESGAOLNELYKYSILHMGHGMIAFLPIINOLFRVLTFRYDQNEVA
VNTVYTIHYVACQHEBGLSLRSTYKAYKAEPIYASEKYTVHEELKSMPTILKP
SADFLSNKLLKYSWFFDYLLIKSMAQHLLENKVKLNRQRPASVYHAIVEVYVNL
MPHTOKFRDNPSPASKNANSHLAVETIKRCPTEGDFVFKOINNYISCPADPKYLE
EYKFEFLRVYCNHEHYIPLNPMFPGKGRFORODLOLQSLTDECEHNLFLYGLLR
EYGTALOEFEVRLIAISYKLLIKHSFDRASRASHORALITLYPFGLLIEVVO
RINVRDYSPPRVNAGMTVRDESLALPVRNLTVPKQGSTLNDLSHLDGLISATSP
YTTSTPINSVNRANDSRGSLISTDSGNSLPERSEKSNISLDKQOSTIGNSVYRDK
LDQSEIKSLMCELYLILKMSDPALEFTYMKASTSLMDPFTISEVCLHQFOYMGKRY
IARTGMHARLQOLGSLDNLSPFNHSGHSDADVLQSLLEANIAVECLTALDLSL
FTLAKNQLADHGHNLMAKVPDYLCFLQKHQSTALKNFTALRSILYKFSPEY
EGRADMCALCYELKCCNSKLSIRPEASQLLYFLMRNNPDYTGKSFVRHLQVYI
SVSLTADVYIGGTRPOGSLTINNCANDRLIKHRSSDPKDLFRKLRYLMAFA
QMKEHENDPMLVDLOTSLKASTIPDELKRTLDSMARLHYKNGLSERACVYVNT
ALVAEYLTREAVQWEPPLLPHSHSACLRSRGVGRQGTAFRVTTPNIDERSME
DVGMQDYHFNEVDYLMELLEQCADGLMAERYELIADYIKLIPIYERKDFERLHLY
DTLHRASKTYEVNHSGRLLGTFRFAFGQAAYOFTDSEFDVGEFDEDEGKEYI
YKEPKLPLSEISORILKILSDKRGSNVMIQDSGVNKRDLDSKYAAIYOVHVIYF
FDEKELOERKTERPSHNRIRREMEFMTQYRQGGVEQCKRRTILTAIHYCFYKA
KRIVVYQHTDLPILVAILDEMSKYAEIROLICSSAEFDMIKLOIKLOGSVSYOVA
GPLVAAPAFILDTNTRYPDNKVKLKEYVRQFEVACGQALAVNERLIKEDOLEYOE
MKANVRMAKELSEIMHEOLG"

BASE COUNT 2175 a 1687 c 1753 g 1930 t
ORIGIN

Query Match 99.7%; Score 498.4; DB 9; Length 7545;
Best Local Similarity 99.8%; Pred. No. 3,7e-146;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTTTACACCATCCCAAAACCCGAAATTTATGATGATTAATAATAGAGTGGCCAC 60
|||||
Db 2178 AGTTTACACCATCCCAAAACCCGAAATTTATGATGATTAATAATAGAGTGGCCAC 2257
QY 61 TCAGCTGATGAAAGACCAACCTGTTGTCACATTTCCATGTCAGCTGCAACTC 120
|||||
Db 2238 TCAGCTGATGAAAGACCAACCTGTTGTCACATTTCCATGTCAGCTGCAACTC 2297
QY 121 AAGTAAAGGAAGCAGAGGAAGAGGATGTCGTGAACCCCAAGTGGCTCTCCTGCT 180
|||||
Db 2298 AAGTAAAGGAAGCAGAGGAAGAGGATGTCGTGAACCCCAAGTGGCTCTCCTGCT 2357
QY 181 TCCCTCCTGAAAGAGGAGGGTGGTGAAGAAGAGGACACATCCCGGCTCGGCGAA 240
|||||
Db 2358 TCCCTCCTGAAAGAGGAGGGTGGTGAAGAAGAGGACACATCCCGGCTCGGCGAA 2417
QY 241 CCTTCCTTGGGGCTATCTTGGCTACCAAGAGCTGGGATGGGACAGCATTAATGCTCCGGA 300
|||||
Db 2418 CCTTCCTTGGGGCTATCTTGGCTACCAAGAGCTGGGATGGGACAGCATTAATGCTCCGGA 2477
QY 301 AATTAAATGGGTAGATGAGGCAAGCCACTGCTGAAATTTCCACATCTGTTTCTAC 360
|||||
Db 2478 AATTAAATGGGTAGATGAGGCAAGCCACTGCTGAAATTTCCACATCTGTTTCTAC 2537
QY 361 AGTGTATCTACAGATCAGCATTTACATTAATTTTTCAGTACTGTGAGAAACCGAATC 420
|||||
Db 2538 AGTGTATCTACAGATCAGCATTTACATTAATTTTTCAGTACTGTGAGAAACCGAATC 2597
QY 421 TGGAGCCCAAGCCTTGGAAACGAACCTTGAAGTACCTTAAGAGTCGATGATGGA 480
|||||
Db 2598 TGGAGCCCAAGCCTTGGAAACGAACCTTGAAGTACCTTAAGAGTCGATGATGGA 2657
QY 481 AGGCACGTCGATGATGCTCCT 500
|||||
Db 2658 AGGCACGTCGATGATGCTCCT 2677

LOCUS	AX173022	6454 bp	DNA	Linear	PAT 03-JUL-2001				
DEFINITION	Sequence 1 from Patent WO0142294.								
ACCESSION	AX173022								
VERSION	AX173022.1	GI:14597975							
KEYWORDS									
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
AUTHORS	Lu, P., Garman, D.D. and Candia, A.F.								
TITLE	Clasp-4, transmembrane protein								
JOURNAL	Patent: WO 0142294-A, 14-JUN-2001.								
FEATURES	<p>Atdor Vita Corporation (US)</p> <p>Location/Qualifiers</p> <p>1..6454</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/note="full length human Clasp-4 cDNA"</p> <p>95..6121</p> <p>/note="human Clasp-4"</p> <p>/codon_start=1</p> <p>/protein_id="CAC43729.1"</p> <p>/db_xref="GI:14597976"</p> <p>/translation="MFPEDISIVGRVOSTVEDAEKRAQSIFVEKCIKTVS TDMHVVNYKYEDSGDFRMLPKSLRPKEIPVHPVEIDEDCEKEDSSSLCSQKGVIT KQOGLKANKVNSTIYTMKVFYKRRFYITQLPDGSIINSTKDEKNSKSGCIYDIA CIVDVOCPKMRHRAELKAKLNDYSHLAELEQEEHMLITLKLITQINTSLQVEKK EIVEIADODETSQGAENIMASLSERSKAGRPTEQLINKLSRDGRONLFSFD SEVQRIDFSGIEPDIKFPBEKCKRFFLYNCHDLPENIGQIGDNAGKGPVVEPFIY LAEVRIDNNCKISADPHVDLNPSPVREMIMGSTQIADSGSKSSPESTHGAESO LRYIOGIFSVTPNHPPEIFLVARIKVLOGNTHCAEPIKNSDPVTKAQVHRHTAKO VCSRLGQYMPFAMARPIFKDTGSLDLDGSEPLKYKDSKISSEDIKLISYKAK PECTKIOIIRGOLNITVECPVDSLNCITSSVPLKPEKKNQNTIYVEEVPBMTK YCTPIPTIKNLVYPLQKLTDSQTEKAKRIAVCFEROSDSDASLKCIYGRPA GSVFITNAIVVSHNQNEFEYDEIKELPHLQKHNLITFYVSCETIKTKTKK QDVTPEVGFANVPLKDKRIITPEQOLPVSANLPGLYLNIDAESRQCVIDIKWD GAPDLKFSKHESTIYTQDILVHKFEHFCOLIOGSKSEVGEILKYLKCLHAMEIOV MEDELVIIMOLFETVLTNTMHDVPICTMVLHTHYKCHGEGLDILRSFIYSR PEKSPAOAOLIHETLATMTAIIKOSADSEITINKLKYSPFPBIIIAKSATYLLEE NKTKIPROGRPEPTYHHVHLSLLIITHVITTRAIETDESRYNYSIASLKLCITL MDGFLENLINDISGFSKDPKVLAEKFEFLQTCHEHYIPILNLPMAFKRLQD VDQSNLEYSLSDEYCKHPELVGLIRETSLAODNYELRYAIVINLILKAFDTR YQKHOQAQIAQLYLPFVGLILENTOIRAGSDTLYSCAMPNSARDEFPQGFSPAN RGSISTDKDTAAGSPONGHIGIKREDSSSLPEGATGPPODNGENTROSTSSVS QYVRIDPOYEIRSLMKCYLYIKVMSIEDTLILKYSPOELINTLILECPIPERMY GKNRIARVHADVLISKHCGIDRKSQMPILNRSGVMORLQHSLSLSSPTLNHSST TEADITHQALLSGNTATEVSLVLEDTITFEFCFCTQLLNDGNLPMKRVFDHIAF LKKGQSEVSLKHFVAFSLRAFLISKPSAFKGRVNNCAFCYELKCTSKISSTRNA SALLYILMRNNEVYTRKRFELRTHLOITIAVSOILADVALSGSRPOESPIINNEAN SRPMKATAPFAEVNDLTRITRTVYMAAOKMKEHKEDEMIIDLOYLSAKVASTPEL RKWMLDSMAKIHVKNGDSEAMCVYHAAVAELHEKDEKEMIDLOYSGAFKPIPNIDE EGMKREDAGMDVHYSEVYLLELEOCYDGMKAEVYIISLTKLPIYKRRPEE KLQVYRTHGMAVYTKILFEMHKKRLLDTFFRVAVAFQGSFPEEDGKEIYIKBEKLT LSBISIRLKLKIGKFGTEBNKTIIDSDKVAKELDQVYIAHOVYKVPFDEKELTE RKTEFERNNISIEFVEADPYTLISGKQCIQECQKRIITLTSNPSYVAKRIPIEN QOINLEPIDGATDEIKDITAELOKCSSTDVMDIOLQKLOGVSVYVNGRIPIAAE FLUDSGAKYPPKYSLEKDMFRKFIQACSIALENELIKEDOVEHEGKSNFRDM VKLSLDIHQEILQDTHMSPMSMTLHFCALISGTSIDRYSGRARAEV"</p>								
BASE COUNT	2119 a	1234 c	1303 g	1798 t					
ORIGIN									
Query Match	40.9%; Score 204.4; Db 6; Length 6454;								
Best Local Similarity	63.4%; Pred. No. 4.3e-53;								
Matches 313; Conservative	0; Mismatches 181; Indels 0; Gaps 0;								

QY	62	CAGCGCATGTAAAGAACACACACCTGTGGTCTACATTTCTCCATGTCAGTGGACAACTCA	121
Db	2060	CACCTACATCTCAAAAACATCATCTTTGCTTTTCTCTTTTATCATGTGAATTTAAAC	2119
QY	122	AGTAAAGGAGACACAGAGAGAGGATGTCTGTGAAACCCCAAGTTGGCTACTCTGGCTT	181
Db	2120	ACAAAGGGAACACCAAAAACGACAGACACTGTGAACCTCAGTTGGGTTTGCTGGGTA	2179
QY	182	CCCTCTCTGAAAGACGGAAGGGTGTGACAGCAGCAGCATCCGGTCTCGGCGAAC	241
Db	2180	CCTTTGGCTGAAAGATGTGTAGATCATCATCATTTTGGACAGCAGCTCCAGTTCGGCCAA	2239
QY	242	CTTCTTTGGGCTTACTTGTGGCTTACCAAGAGCTTGGGATGGGACAGCATTTATGTCGGAA	301
Db	2240	CTTCCCCAGGCTCTCTTGAATCTCAATGATGCAATTCAGAAAGGCAATGTATACGTGGAT	2299
QY	302	ATTAAATGGGTAGATGTGAGGACGACCACTGCTGAAAAATTTCCATCAATCTGTTTCTACA	361
Db	2300	ATTAAATGGGTAGATGTGAGGACCAACCCCTTTGTGAAAGTTTAAACCACCTTAGATCTAAC	2359
QY	362	GTGTAATCTCAGGATCAGCATTTTACATTAATTTTTCCAGTACTGTACGAAAACCGAATCT	421
Db	2360	ATTACACTCTCAAGATCTGCAATGTGTCACAAATTTCTTCATCATTTGCCAGCATTCAGTCA	2419
QY	422	GGAGCCCAAGCCTTAGAGAAACGACTGTGAAGTACCTTGAAGTCTGCAATGCCATGGAA	481
Db	2420	GGCTGTGAAAGAGATTCAGGGGAGCTCATTTAAATTTTAAAGTGTGTTGATTCATCGATGAG	2479
QY	482	GGCCACGATGAT	495
Db	2480	ATCCAGTCATGAT	2493
RESULT 5			
CNS01RGX	139887 bp	DNA	linear PRI 26-APR-2001
LOCUS	Human chromosome 14		from chromosome 14 of Homo sapiens (Human), complete sequence.
DEFINITION	AL160233		
ACCESSION	AL160233.3	GI:13016589	
VERSION			
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 139887)		
AUTHORS	Heilig,R., Pelt,J.L., Vico,V., Dasilva,C., Robert,C., Winkler,P.,		
	Brolier,P., Cotto,L., Barbe,V., Pelletier,E., Artiguenave,F.,		
	Levy,M., Eckenberg,R., Bruls,T., deBerardins,V., Cruaud,C.,		
	Grayay,G., Saurin,W. and Weissbach,J.		
TITLE	Sequencing of the human chromosome 14		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 139887)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :		
	BP 101 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
	On Feb 20, 2001 this sequence version replaced gi:7799784.		
COMMENT	Genome Center		
	Center: Genoscope / Centre National de Sequencage		
	Center code: GS		
	Web site: http://www.genoscope.cns.fr/		
	Contact: SeqRef@genoscope.cns.fr		

	The following BAC sequence is oriented from the r7 to the SP6 end.		
	Upstream BAC (overlapping the r7 end) : R-857B24 (AC-AL049870)		
	Downstream BAC (overlapping the SP6 end) : R-390E10		
	Summary Statistics		
	Assembly program: Phrap; version 2.0		
	Quality coverage: 8.87x in Q20 bases; sum-of-contigs		

m1.5

```

sequence.
AL161420
AL161420.10 GI:10443397
HTG: KIAA0694; KIAA1058.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 163316)
Smith, M
Direct Submission
Submitted (31-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Oct 1, 2000 this sequence version replaced gi:10039689.
During difference assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
This sequence is the entire insert of clone RP11-155N3 The true
left end of clone RP11-18G11 is at 114983 in this sequence. The
true right end of clone RP11-56D6 is at 42341 in this sequence. The
true right end of clone RP11-551M18 is at 43464 in this sequence.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
RP11-155N3 is from the library RPI1-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pACe3.6.
Location/Qualifiers
1..163316
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-155N3"
/clone_id="RPI1-11.1"
6..510
/note="match: GSS: Em:AQ372034"
75..372
/note="AluYb repeat: matches 1..298 of consensus"
447..742
/note="AluSp repeat: matches 1..296 of consensus"
745..1025
/note="L1Mc1 repeat: matches 6019..6327 of consensus"
1250..1566
/note="AluXx repeat: matches 2..311 of consensus"
1631..2325
/note="L1Mc5 repeat: matches 6683..7400 of consensus"
2326..2619
/note="AluXx repeat: matches 1..295 of consensus"
2620..2890
/note="L1Mc5 repeat: matches 7400..7665 of consensus"
complement(3019..3311)
/note="match: STS: Em:HSA273ZH1"
3060..3191
/note="match: GSS: Em:BI7811"
complement(3065..3326)
/note="match: STS: Em:AF009212"
3067..3190

```

```
repeat_region /note="match: GSS: Em:AQ114323"
3087. .3122
/note="18 copies 2 mer tg 97% conserved"
repeat_region 3089. .3124
/note="6 copies 6 mer tgggtg 97% conserved"
misc_feature 3123. .3280
/note="match: GSS: Em:AQ459714"
3191. .3475
/note="Alusx repeat: matches 1. .288 of consensus"
repeat_region 3692. .3773
/note="MIR repeat: matches 3. .94 of consensus"
3887. .4074
/note="L2 repeat: matches 1607. .1787 of consensus"
repeat_region 4082. .4144
/note="3 copies 21 mer 88% conserved"
4418. .4630
/note="L2 repeat: matches 1354. .1587 of consensus"
7208. .7609
/note="MIR repeat: matches 8. .390 of consensus"
9220. .9344
/note="MIR repeat: matches 71. .196 of consensus"
9586. .9667
/note="L1ME repeat: matches 5569. .5651 of consensus"
9714. .9813
/note="L2 repeat: matches 771. .874 of consensus"
9869. .10202
/note="L2 repeat: matches 971. .1261 of consensus"
10203. .10581
/note="MIRA repeat: matches 1. .426 of consensus"
10582. .11267
/note="L2 repeat: matches 1261. .2121 of consensus"
11460. .11570
/note="L2 repeat: matches 2377. .2497 of consensus"
11603. .11778
/note="L1MB3 repeat: matches 5996. .6166 of consensus"
11779. .12160
/note="MER51A repeat: matches 1. .371 of consensus"
12161. .12729
/note="L1MB3 repeat: matches 5403. .5996 of consensus"
12682. .12950
/note="L1MD repeat: matches -1. .274 of consensus"
12945. .13209
/note="L1MCS repeat: matches 7249. .7517 of consensus"
13323. .13413
/note="Alu/FRAM repeat: matches 221. .311 of consensus"
13711. .13846
/note="Alusg/x repeat: matches 1. .136 of consensus"
13955. .14417
/note="match: GSS: Em:AQ475161"
complement(14122. .14590)
/note="match: GSS: Em:AQ349540"
14334. .14400
/note="L1MA10 repeat: matches 6204. .6272 of consensus"
14407. .14693
/note="Alusx repeat: matches 1. .288 of consensus"
14584. .15413
/note="match: GSS: Em:AQ739876"
14719. .14793
/note="L1MC/D repeat: matches 5744. .5818 of consensus"
15013. .15193
/note="L1MB3 repeat: matches 5993. .6184 of consensus"
complement(join(16480. .17741,19199. .19206,20108. .20246,
20382. .20482,23281. .23471,27992. .28202,30654. .30800,
31633. .31668,32116. .32121,32345. .32458,33162. .33284,
47391. .47497,48865. .48933,49823. .49880,52300. .52515,
52636. .52749,54341. .54523,54642. .54779,60487. .60602,
68322. .68365,68917. .69053,71585. .71632,73040. .73106,
76401. .76515,78891. .79024,83174. .83515,86012. .86125,
86363. .86541,90418. .90529,90921. .91004,91224. .91370,
102891. .102962,103540. .103665,104551. .104620,
104885. .105000,106035. .106124,106806. .106892,
107964. .108085,108701. .108824,109525. .109617,
111151. .111216,111353. .111540,111625. .111704,
120481. .120603))
/gene="bA155N3.2"
mRNA
complement(join(16480. .17741,19199. .19206,20108. .20246,
20382. .20482,23281. .23471,27992. .28202,30654. .30800,
31633. .31668,32116. .32121,32345. .32458,33162. .33284,
47391. .47497,48865. .48933,49823. .49880,52300. .52515,
52636. .52749,54341. .54523,54642. .54779,60487. .60602,
68322. .68365,68917. .69053,71585. .71632,73040. .73106,
76401. .76515,78891. .79024,83174. .83515,86012. .86125,
86363. .86541,90418. .90529,90921. .91004,91224. .91370,
102891. .102962,103540. .103665,104551. .104620,
104885. .105000,106035. .106124,106806. .106892,
107964. .108085,108701. .108824,109525. .109617,
111151. .111216,111353. .111540,111625. .111704,120481. .>120603))
/gene="bA155N3.2"
/product="bA155N3.2.1 (K1AA10581)"
/note="match: cDNAs: Em:AB028981 Em:U79289
match: ESTs: Em:W27872 Em:H29845 Em:H29936 Em:AW160928
Em:AW457794 Em:T62004 Em:R64450 Em:D58813 Em:H10596
Em:AW480042 Em:AW344638 Em:AW450751 Em:AF188523
Em:AA151055 Em:T33377 Em:AW466130 Em:AI014329 Em:Z42960
Em:AI372817 Em:AI372818 Em:AW389763 Em:AW230012
Em:AI133724 Em:AW162535 Em:AW607903 Em:AA484945 Em:H28862
Em:Z39079 Em:AI67696 Em:AW576807 Em:AI078244 Em:R78583
Em:AA705585 Em:AI271453 Em:AW512964 Em:W88550 Em:AI870702
Em:AL079478 Em:AI693777 Em:AW576805 Em:AI467594
Em:AI172819 Em:AI732576 Em:AI277278 Em:AI69485
Em:AI768094 Em:AA779769 Em:AA502043 Em:R27124 Em:AW576813
Em:AW576816 Em:AW168032 Em:AI165152 Em:AA165512 Em:AA523853
Em:AI120344 Em:AI888868 Em:AA165554 Em:R73986 Em:AA523853
Em:AI041815 Em:AL134605 Em:AW009821 Em:D62808 Em:AI769357
Em:AW363093 Em:AI733873 Em:W88496 Em:T03019 Em:AW071481
Em:AI034000"
/evidence="not_experimental"
polyA_site complement(16480)
/gene="bA155N3.2"
polyA_site complement(16482)
/gene="bA155N3.2"
polyA_site complement(16483)
/gene="bA155N3.2"
polyA_signal complement(16496. .16503)
/gene="bA155N3.2"
misc_feature complement(16523. .16528)
/note="match: STS: Em:G42960"
polyA_signal complement(16523. .16528)
/gene="bA155N3.2"
polyA_signal complement(16570. .16575)
/gene="bA155N3.2"
complement(join(17741,19199. .19206,20108. .20246,
20382. .20482,23281. .23471,27992. .28202,30654. .30800,
120481. .120603))
Query Match 25.4% Score 126.8; DB 9; Length 163316;
Best Local Similarity 98.5% Pred. No. 2.9e-28;
Matches 128; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 39 AGATTAAATAGACTGTGCCACTCAGCTGCATGTAAGACACACCTGTCATATTC 98
|||||
Db 108826 AGATTAAATAGACTGTGCCACTCAGCTGCATGTAAGACACACCTGTCATATTC 108767
|||||
QY 99 TCATGTGACGTGTGACACTCAAGTAAGAGACAGACAGAGAGGATGTCTTAAA 158
|||||
Db 108766 TCATGTGACGTGTGACACTCAAGTAAGAGACAGACAGAGAGGATGTCTTAAA 108707
|||||
QY 159 CCCAAGTGG 168
|||||
Db 108706 CCCAAGTGG 108697
|||||
RESULT 7
AC109966/c AC109966 175281 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-33315. *** SEQUENCING IN PROGRESS
DEFINITION *** 75 unordered pieces.
```

```
ACCESSION      AC109966
VERSION        AC109966.3
KEYWORDS       GI:21738210
SOURCE         HTG: HTGS PHASE1.
ORGANISM       Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 175281)
AUTHORS        Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Aisbrook,S.L., Amaralunge,H.C., Are,D.R., Ayele,M., Banks,T.,
Barbada,J., Benton,J., Bimaga,K., Blankenburg,K., Bonin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,D.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsl,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,R.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Moser,M., Neal,D., Newton,S., Newson,E., Nwokenkwo,S., Oguh,M., Morris,S.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunnu,G.,
Otagunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sutton,G., Sonake,T., Sparks,A., Stanley,H., Stone,K.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansay,U., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,O.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
TITLE          Direct Submission
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 175281)
AUTHORS        Worley,K.C.
TITLE          Direct Submission
JOURNAL        Submitted (09-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 175281)
REFERENCE      Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18847026.
JOURNAL        ----- Genome Center
AUTHORS        Center: Baylor College of Medicine
TITLE          Center code: BCM
JOURNAL        Web site: http://www.hgsc.bcm.tmc.edu/
COMMENT        Contact: hgsc-help@bcm.tmc.edu
Project Information
Center Project name: GOLF
Center Clone name: CH230-33315
----- Summary Statistics
Sequencing vector: Plasmid.
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 105442 bases at least Q40
Consensus quality: 111492 bases at least Q30
Consensus quality: 116270 bases at least Q20
-----
NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 75 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
1291 1290: contig of 1290 bp in length
1391 1390: gap of unknown length
1391 2414: contig of 1024 bp in length
2415 2514: gap of unknown length
2515 3639: contig of 1125 bp in length
3640 3739: gap of unknown length
3740 5063: contig of 1324 bp in length
5064 5163: gap of unknown length
5164 6243: contig of 1080 bp in length
6244 6343: gap of unknown length
6344 7336: contig of 1193 bp in length
7337 9071: contig of 1435 bp in length
9072 9172: gap of unknown length
9172 10200: contig of 1029 bp in length
10201 10300: gap of unknown length
10301 11505: contig of 1205 bp in length
11506 11605: gap of unknown length
11606 12680: contig of 1075 bp in length
12681 12780: gap of unknown length
12781 14056: contig of 1276 bp in length
14057 14156: gap of unknown length
14157 15413: contig of 1257 bp in length
15414 15513: gap of unknown length
15514 16580: contig of 1067 bp in length
16581 17692: gap of unknown length
17692 17792: contig of 1012 bp in length
17793 19013: contig of 1221 bp in length
19014 20418: contig of 1305 bp in length
20419 20518: gap of unknown length
20519 21880: contig of 1362 bp in length
21881 21980: gap of unknown length
21981 22985: contig of 1005 bp in length
22986 23085: gap of unknown length
23086 24156: contig of 1071 bp in length
24157 24256: gap of unknown length
24257 25732: contig of 1476 bp in length
25733 25832: gap of unknown length
25833 27196: contig of 1364 bp in length
27197 27296: gap of unknown length
27297 28615: contig of 1319 bp in length
28616 28715: gap of unknown length
28716 30145: contig of 1430 bp in length
30146 30245: gap of unknown length
30246 31418: contig of 1173 bp in length
31419 31518: gap of unknown length
31519 32896: contig of 1378 bp in length
32897 32996: gap of unknown length
32997 34142: contig of 1146 bp in length
34143 34242: gap of unknown length
34243 35262: contig of 1020 bp in length
35263 35362: gap of unknown length
35363 36924: contig of 1562 bp in length
36925 37024: gap of unknown length
```

	*	37025	38670:	contig of 1646 bp in length
	*	38671	38770:	gap of unknown length
	*	38771	40612:	contig of 1842 bp in length
	*	40613	40712:	gap of unknown length
	*	40713	42724:	contig of 2012 bp in length
	*	42725	42824:	gap of unknown length
	*	42825	43829:	contig of 1005 bp in length
	*	43830	43929:	gap of unknown length
	*	43930	45818:	contig of 1889 bp in length
	*	45819	45918:	gap of unknown length
	*	45919	48183:	contig of 2265 bp in length
	*	48184	48283:	gap of unknown length
	*	48284	49887:	contig of 1604 bp in length
	*	49888	49987:	gap of unknown length
	*	49988	51568:	contig of 1581 bp in length
	*	51569	51668:	gap of unknown length
	*	51669	53251:	contig of 1583 bp in length
	*	53252	53351:	gap of unknown length
	*	53352	55636:	contig of 2285 bp in length
	*	55637	55736:	gap of unknown length
	*	55737	57159:	contig of 1423 bp in length
	*	57160	57259:	gap of unknown length
	*	57260	59519:	contig of 2260 bp in length
	*	59520	59619:	gap of unknown length
	*	59620	63392:	contig of 3773 bp in length
	*	63393	63492:	gap of unknown length
	*	63493	65130:	contig of 1638 bp in length
	*	65131	65230:	gap of unknown length
	*	65231	68145:	contig of 2915 bp in length
	*	68146	70070:	contig of 1825 bp in length
	*	70071	70170:	gap of unknown length
	*	70171	73494:	contig of 3324 bp in length
	*	73495	73594:	gap of unknown length
	*	73595	75991:	contig of 2397 bp in length
	*	75992	76091:	gap of unknown length
	*	76092	77867:	contig of 1776 bp in length
	*	77868	77967:	gap of unknown length
	*	77968	80318:	contig of 2351 bp in length
	*	80319	80418:	gap of unknown length
	*	80419	82211:	contig of 1793 bp in length
	*	82212	82312:	gap of unknown length
	*	82312	84412:	contig of 2101 bp in length
	*	84413	84512:	gap of unknown length
	*	84513	85982:	contig of 1470 bp in length
	*	85983	86082:	gap of unknown length
	*	86083	88348:	contig of 2266 bp in length
	*	88349	90746:	contig of 2298 bp in length
	*	90747	90846:	gap of unknown length
	*	90847	92600:	contig of 1754 bp in length
Query Match				
Best Local Similarity 83.1%; Score 90.4; DB 2; Length 175281;				
Matches 103; Conservative 0; Mismatches 21; Indels 0; Gaps 0;				
Qy	163	AGTTGGCTACTCCTGGCTTCCCTCTGAAAGAGCGAAGGCTGTACACAGCAGCAGCA	222	
Db	84981	AGTTGGCTTCTCCTGGCTTCTCTGAAAGAGCGAGAGGTTGACGATGTAACAGCA	84922	
Qy	223	CATCCCGGCTCGGCGAACCCTTCTCGGCGATCTTGGCTACAGAGAGTTGGATGGG	282	
Db	84921	TATCCCTGTCTCCGCCACCTGCGCTGTGCTACCTTGGCTACAGAGAGCTCAGCATGGG	84862	
Qy	283	CAGG 286		
Db	84861	CAGG 84858		
RESULT 8				
LOCUS AX173028 4391 bp DNA linear PAT 03-JUL-2001				
DEFINITION Sequence 7 from Patent WO0142294.				

ACCESSION AX173028				
VERSION AX173028.1 GI:14597978				
KEYWORDS				
SOURCE human.				
ORGANISM Homo sapiens				
REFERENCE 1 (bases 1 to 4391)				
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.				
TITLE Clasp-4 transmembrane protein				
JOURNAL Patent: WO 0142294-A 7 14-JUN-2001;				
Arbor Vita Corporation (US)				
FEATURES				
source Location/Qualifiers				
1..4391				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/note="human CLASP-4 cDNA"				
414..4058				
/note="human CLASP-4"				
/codon_start=1				
/protein_id="CAC43730.1"				
/db_xref="GI:14597978"				
/translation="MEIQVLIRFLSVILMQLFVLEPNMIEDDVPISCPWLFHYSK				
CHEGLDYSLSFRIKYSFRGKPSAPQAPLHETLATMMIALKQSADELAINKLKY				
SWFEFEIYAKSMATVYLEENKIKLTHGQRPKYLHSLFLAIFVESQVALEPKR				
SRNNTSLASELKCCLTMDRGVFNILINDYISGFSPKDPKFLAEKFELOICNHE				
HYPLINPMAFPAPKIQROVDSNLEYSLSDEYKHFHVLGILIRFSTALODNYEIRY				
TALSVINLILIKIAFDTRIYHKNQAKIAOLYLPFGLLLENIRORLAGRDITSCAM				
PNSNRDEPFGFTSPANRGLSTPDQTAGSFQNGHGKREDROSLLPEGATGPD				
QNGMNGENTROSTRSSVQYRNLDQYEIRSLMCYLYIKVMSEDLTYWNKVSPOE				
LINLILILEVCLFHFRRMGKRNIARVHADVAKSHFGIDRSQTMPLRRSGVQARL				
OHLSLSESTLNHSSSTTEADIEHQLLEGNATVSLVDLTIEFFQCFETFLN				
NDGNIPLMKRVPHIHLAEIKNGOSEVSKIVSLKAFISKPSPAFKRVNCAACF				
YELVLCCTKSTISTRNDSALYLILHRNNEFYRKRTFLRTHQIIITANISOLIDVAL				
SGGRFQESFLLINNEFNSDRPMILARFPAEVADLYTRKIRTVLMATQKHEKDEEM				
LIPDQSLASVASTPELRTKWLDSMAKIHVKNQDSEAMCYVAALVAERLHKR				
LEPFGCSAFKRTIPNIDEGAMKEDGMVDVHSEVYLELEOCYVNGMKAEYELI				
SEISKILGPIYENREFENLOUYRTLHGAYTKILFMYHTRKRLGLTFPRFAVGGSF				
FEEDDEKEVLYKRPKLGISELIRLVKLYGKEGEMVNYKIIODDSKVAKLEDPYA				
HIQVTTYKPIPFDDKELTERKTEFERNNHNSRFEAPRYTSLGKQOCIEQCKRRTIL				
TNTSFPYVKKRIPIETNCEQIINKPIDGATDEIKDKTAELQKLCSTTDMDIOLKL				
QGWTSVQVNAKPIYARAFNDQSASVYPRKVSLEKDMFRKTIQACSIALLENELI				
KEDQVEHEGLKISFRDMYKELSDIIHQIQLDPTMHPMSMTLHVFCALISGTSDDR				
GYGSPRYAEV"				
BASE COUNT 1359 a 851 c 908 g 1273 t				
ORIGIN				
Query Match				
Best Local Similarity 17.0%; Score 85; DB 6; Length 4391;				
Matches 187; Conservative 0; Mismatches 170; Indels 0; Gaps 0;				
Qy	140	AAGAGGATGTCGTTGAACAACCAAGTTGGCTACTCCTGCTCCCTCTCGAAGACGCA	199	
Db	78	AAGGCGGCCCCGGTGGACCTCCCGGTGGGTTGGCTGGCTTCGTAAGGTTGGT	137	
Qy	200	AGGCTGTGACAAAGGACACACATCCCGGTCTCGGCGAATCTTCTCGGGCTATCTT	259	
Db	138	AGATTCTTCCTTTTGGGCGGCGCTGCCGTTTCCCATCTTCCCGCGGCTCTCTTA	197	
Qy	260	GGCTACCAAGACCTTGGATGGGACGATTATGCTCCGGAATTAATTAATGATGATGA	319	
Db	198	ATTCTGATTGATGATTAATCAAGGGGCAATGATATGGAATTTAAATGGGTGGTGT	257	
Qy	320	GGCAAGCCACAGCTCTAAATTTCCACTATGCTGGTTTCTACAGTGTATACACAGATCAG	379	
Db	258	GTTAAAGCCTTGTGTGAGGTTTAAAGCCCCCTTGTCTCCCATTTCCCTTAAGGTTCTG	317	
Qy	380	CATTACATTAATTTTCTCCAGTACTGCAGAAACCGAATCTGAGGCCACAGCCTTAGGA	439	
Db	318	CATTGGCCCAATTTCTTTCATCATATTCGCGCTGATGTGGTACAGCTCGAAGAGGTTTCCA	377	
Qy	440	AACGACTTGTAACTACCTTAAGAGTCTGCATGGATGGAAGGCCACGTGATATC	496	

DB	378	GGGGGGGCTCATTAATTTTAAAGCTCTTGATGCGCATGAGAGATCCAGATCTTGATC	434
RESULT 9			
LOCUS	AX173118	4393 bp	Linear
DEFINITION	Sequence 97 from Patent WO0142294.		
ACCESSION	AX173118		
VERSION	AX173118.1	GI:14598012	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Source			
BASE COUNT	1359 a	852 c	910 g 1272 t
ORIGIN			
Query Match	17.0%;	Score 85;	DB 6; Length 4393;
Best Local Similarity	52.4%;	Pred. No. 2.5e-15;	
Matches 187;	Conservative 0;	Mismatches 170;	Indels 0; Gaps 0;
QY	140	AAGAGGATGTCGTGGAACCCCAAGTGGCTACTCTCGCTCCCTCTCGTAAGACGGA	199
DB	78	AAGGGGGCCCCGGGTGACCTCCGGTTGGCTGGCTGGCTCCCTTCGTGAAGGTGGT	137
QY	200	AGGGTGGTACAAAGGACGACGACATCCCGGCTCGGCGAACCCTTCCTGGAGTATCTT	259
DB	138	AGATCTTCCCTCTTTGGGCGGCGCTGCGCGTTTCCCATTTCTTCCCGGCTCCTTA	197
QY	260	GCGTACCAAGACCTTGGAGTGGCGACGATTAATGCTCCGGAATTAATGGCTAGATGGA	319
DB	198	ATTCTGATTTGATGATGATTAATCAAGGGGCAATGTATGATTTAAATGGGTGGTGGT	257
QY	320	GCGAGCGACGTCGTAAATTTCCACATCATCGGTTTCTACAGTGTATACATAGATCAG	379
DB	258	GTAAGCCCTTGGAGGTTTAAAGGCCCTTGGGTTCTCCATTTCCCTTAGGTTGCG	317
QY	380	CATTACATTAATTTTTCACGATGTCGAGAAACCAATCTGAGCCCAAGCCTTAGGA	439
DB	318	CATTGGCCCAATTTCTTCATCATTTGCGCGCTGATTTGCGTCAAGTCCGAAGAGTTCCA	377
QY	440	AACGAACCTTGAAGACTACCTTAAGAGTGTGCATGCGATGGAAGGCCAGGTGATGATC	496
DB	378	GGGGGGGCTCATTAATTTTAAAGTCTTTGATGCGCATGAGAGATCCAGATCTTGATC	434
RESULT 10			
LOCUS	AC109966	175281 bp	DNA
DEFINITION	Rattus norvegicus clone CH230-32315, *** SEQUENCING IN PROGRESS		
ACCESSION	AC109966		
VERSION	AC109966.3	GI:21738210	
KEYWORDS	HTG; HTGS; PHASE1.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 175281)		
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Gosman,F.R., Allen,C.,		
	Alshrocks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,		
	Barbataia,J., Benton,J., Blumage,K., Blankenburg,K., Bonnin,D.,		
	Boucek,J., Bowler,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,		

[illegible]

Center Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GOLP
Center clone name: CH230-32315

Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 105442 bases at least Q40
Consensus quality: 111492 bases at least Q30
Consensus quality: 116270 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraft.data.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 75 contigs. The true order of the pieces
* is not known and their order in this sequence reported as
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1290:	contig of 1290 bp in length
1291	1330:	gap of unknown length
1391	2414:	contig of 1024 bp in length
2415	2514:	gap of unknown length
2515	3699:	contig of 1125 bp in length
3640	3739:	gap of unknown length
3740	5063:	contig of 1234 bp in length
5064	5163:	gap of unknown length
5164	6243:	contig of 1080 bp in length
6244	6343:	gap of unknown length
6344	7536:	contig of 1193 bp in length
7537	7636:	gap of unknown length
7637	9071:	contig of 1435 bp in length
9072	9171:	gap of unknown length
9172	10200:	contig of 1029 bp in length
10201	10300:	gap of unknown length
10301	11505:	contig of 1205 bp in length
11506	11605:	gap of unknown length
11606	12680:	contig of 1075 bp in length
12681	12780:	gap of unknown length
12781	14056:	contig of 1276 bp in length
14057	14156:	gap of unknown length
14157	15413:	contig of 1257 bp in length
15414	15513:	gap of unknown length
15514	16580:	contig of 1067 bp in length
16581	16680:	gap of unknown length
16681	17592:	contig of 1012 bp in length
17593	17792:	gap of unknown length
17793	19013:	contig of 1221 bp in length
19014	19113:	gap of unknown length
19114	20418:	contig of 1305 bp in length
20419	20518:	gap of unknown length
20519	21880:	contig of 1362 bp in length
21881	21980:	gap of unknown length
21981	22885:	contig of 1005 bp in length
22886	23085:	gap of unknown length
23086	24156:	contig of 1071 bp in length
24157	24256:	gap of unknown length
24257	25732:	contig of 1476 bp in length
25733	25832:	gap of unknown length
25833	27196:	contig of 1364 bp in length
27197	27296:	gap of unknown length
27297	28615:	contig of 1319 bp in length
28616	28715:	gap of unknown length
28716	30145:	contig of 1430 bp in length
30146	30245:	gap of unknown length
30246	31418:	contig of 1173 bp in length
31419	31518:	gap of unknown length
31519	32896:	contig of 1378 bp in length
32897	32996:	gap of unknown length
32997	34142:	contig of 1146 bp in length
34143	34242:	gap of unknown length
34243	35562:	contig of 1020 bp in length
35563	35624:	gap of unknown length
35625	36824:	contig of 1362 bp in length
36825	37024:	gap of unknown length
37025	38670:	contig of 1646 bp in length
38671	38770:	gap of unknown length
38771	40612:	contig of 1842 bp in length
40613	40712:	gap of unknown length
40713	42724:	contig of 2012 bp in length
42725	42824:	gap of unknown length
42825	43829:	contig of 1005 bp in length
43829	43929:	gap of unknown length
43930	45818:	contig of 1889 bp in length
45819	48183:	gap of unknown length
48184	48283:	contig of 2265 bp in length
48284	49887:	gap of unknown length
49887	contig of 1604 bp in length	

*	49688	49987: gap of unknown length
*	49988	51568: contig of 1561 bp in length
*	51568	gap of unknown length
*	51659	53231: contig of 1563 bp in length
*	53252	53351: gap of unknown length
*	53352	55636: contig of 2285 bp in length
*	55637	55736: gap of unknown length
*	55737	57159: contig of 1423 bp in length
*	57160	57259: gap of unknown length
*	57260	59519: contig of 2260 bp in length
*	59520	59619: gap of unknown length
*	59620	63392: contig of 3773 bp in length
*	63393	63492: gap of unknown length
*	63493	65330: contig of 1658 bp in length
*	65331	65230: gap of unknown length
*	65331	68145: contig of 2915 bp in length
*	68146	68245: gap of unknown length
*	68246	70070: contig of 1825 bp in length
*	70071	70170: gap of unknown length
*	70171	73494: contig of 3324 bp in length
*	73495	73594: gap of unknown length
*	73595	75921: contig of 2397 bp in length
*	75922	76091: gap of unknown length
*	76092	77687: contig of 1776 bp in length
*	77688	77967: gap of unknown length
*	77968	80318: contig of 2351 bp in length
*	80319	80418: gap of unknown length
*	80419	82211: contig of 1793 bp in length
*	82212	82311: gap of unknown length
*	82312	84412: contig of 2101 bp in length
*	84413	84512: gap of unknown length
*	84513	85982: contig of 1470 bp in length
*	85983	86082: gap of unknown length
*	86083	88348: contig of 2266 bp in length
*	88349	88448: gap of unknown length
*	88449	90746: contig of 2298 bp in length
*	90747	90846: gap of unknown length
*	90847	92600: contig of 1754 bp in length

Query Match 15.2%; Score 75.8; DB 2; Length 175281;
Best Local Similarity 68.9%; Pred. No. 4.1e-12;
Matches 104; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY	323	AAGCAGCTGCGTAATTTCCACACTCATCTGGTTCCTAACAGTGTAATCAAGATCAAGCAT	382
Db	83560	AAGCACACATCAACAATTTCCTCTCCTCTCCGCCCTCAACCTGGATCAGCAC	83619
OY	383	TTACATAATTTTTTCAGTACTGTGASAAAACCAGATCTGSAGGCCAAGCCTTAGGAAAC	442
Db	83620	TTACACAAATTTTTTCCAATACGTGCAGAAAACGGAATCTGGAGCCCAAGCCTTAGGGAGT	83679
OY	443	GAACTGTGAAGTACTCTAAGAGTGCATG	473
Db	83680	GAACTAGTGAATACCTTAAGGTACCAAATG	83710

RESULT	11				
LOCUS	AC119357/c				
DEFINITION		192825 bp	DNA	linear	HTG 18-JUL-2002
ACCESSION	AC119357				
VERSION	AC119357.3	GI:21746916			
KEYWORDS	HTG; HTGS_PHASE1.				
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 192825)				
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimode,K., Blankenburg,K., Bonin,D.,				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
	***, 89 unordered pieces.				
	*** SEQUENCING IN PROGRESS				

Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carrott, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 DeLaune, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Girelli, O.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, R., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
 Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratoch, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
 Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Moser, M., Neal, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunnu, G.,
 Oragunye, N., Oyiedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G.,
 Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
 Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansy, U., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, O.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.

Direct Submission
 2 (bases 1 to 192825)
 Morley, K.C.

Direct Submission
 Submitted (26-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 192825)
 Morley, K.C.

Direct Submission
 Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 14, 2002 this sequence version replaced gi:20429771.

Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GVFP
 Center clone name: CH230-473M19

Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 117932 bases at least Q40
 Consensus quality: 124308 bases at least Q30
 Consensus quality: 129203 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 89 contigs. The true order of the pieces
 is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	1129:	contig of 1129 bp in length
*	1229:	gap of unknown length
*	1230	2849: contig of 1620 bp in length
*	2850	2949: gap of unknown length
*	2950	4125: contig of 1176 bp in length
*	4126	4225: gap of unknown length
*	4226	5644: contig of 1419 bp in length
*	5645	5744: gap of unknown length
*	5745	6997: contig of 1253 bp in length
*	6998	7097: gap of unknown length
*	7098	8757: contig of 1660 bp in length
*	8758	8857: gap of unknown length
*	8858	9860: contig of 1003 bp in length
*	9861	9960: gap of unknown length
*	9961	11605: contig of 1645 bp in length
*	11606	11705: gap of unknown length
*	11706	12751: contig of 1046 bp in length
*	12752	12851: gap of unknown length
*	12852	14490: contig of 1639 bp in length
*	14491	14590: gap of unknown length
*	14591	16101: contig of 1511 bp in length
*	16102	16201: gap of unknown length
*	16202	17728: contig of 1427 bp in length
*	17729	18883: contig of 1155 bp in length
*	18884	18983: gap of unknown length
*	18984	20911: contig of 1228 bp in length
*	20911	20911: gap of unknown length
*	20912	21353: contig of 1042 bp in length
*	21354	22518: contig of 1065 bp in length
*	22519	22618: gap of unknown length
*	22619	24107: contig of 1489 bp in length
*	24108	24207: gap of unknown length
*	24208	25603: contig of 1396 bp in length
*	25604	25703: gap of unknown length
*	25704	27171: contig of 1468 bp in length
*	27172	27271: gap of unknown length
*	27272	28993: contig of 1722 bp in length
*	28994	29093: gap of unknown length
*	29094	30122: contig of 1029 bp in length
*	30123	30222: gap of unknown length
*	30223	31317: contig of 1095 bp in length
*	31318	31417: gap of unknown length
*	31418	33144: contig of 1727 bp in length
*	33145	33244: gap of unknown length
*	33245	34431: contig of 1187 bp in length
*	34432	34531: gap of unknown length
*	34532	35743: contig of 1212 bp in length
*	35744	35843: gap of unknown length
*	35844	37180: contig of 1337 bp in length
*	37181	37280: gap of unknown length
*	37281	38596: contig of 1316 bp in length
*	38597	38696: gap of unknown length
*	38697	40031: contig of 1335 bp in length
*	40032	40131: gap of unknown length
*	40132	41485: contig of 1354 bp in length
*	41486	41585: gap of unknown length
*	41586	42744: contig of 1159 bp in length
*	42745	42844: gap of unknown length
*	42845	45158: contig of 2314 bp in length
*	45159	45258: gap of unknown length
*	45259	46752: contig of 1494 bp in length
*	46753	46852: gap of unknown length
*	46853	48029: contig of 1177 bp in length
*	48030	48129: gap of unknown length
*	48130	49682: contig of 1553 bp in length
*	49683	49782: gap of unknown length

Query Match		15.2%: Score 75.8; DB 2; Length 192825;
Best Local Similarity		68.9%: Pred. No. 4.2e-12;
Matches 104; Conservative		0; Mismatches 47; Indels 0; Gaps 0;
Qy	323 AAGCAGCTCTGTAATAATTCACATCTGTTTACAGTATATACAGATCAGAT 382	
Db	33590 AAGCAGCAGCTAAGATATTTCTCTCTCTCCCTGCCCCCTCAACCTAGATCAGCAC 33531	
Qy	383 TTACATATTTTTCAGATCTGACGAAACGGAATCTGGAGCCAGCCTTAGAAGAC 442	
Db	33530 TTACACAAATTTTTCACATATCTGACGAAACGGAATCTGGAGCCAGCCTTAGAGGAGT 33471	
Qy	443 GAACCTGTAAGTACCTTAAGAGTCGCATG 473	
Db	33470 GAACCTAGTGAATACCTTAAGTACCAATG 33440	
RESULT 12		
AK074081	5564 bp	mRNA
LOCUS	Homo sapiens mRNA for FLJ00152 protein.	
DEFINITION	AK074081	
ACCESSION	AK074081.1 GI:18676509	
VERSION	f1s (full insert sequence).	
KEYWORDS	Homo sapiens adult spleen cDNA to mRNA, clone:FLJ00152.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
REFERENCE	1 Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.	
AUTHORS	The nucleotide sequence of a long cDNA clone isolated from human spleen	
TITLE	Published Only in Database (2002)	
JOURNAL		

REFERENCE		2 (bases 1 to 5564)
AUTHORS		Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
TITLE		Direct Submission
JOURNAL		Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3, Yama, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913, Fax:81-438-52-3914)
COMMENT		NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert and 5'- & 3'-end open pass sequencing; Research Association for Biotechnology; cDNA library construction and clone selection; Kazusa DNA Research Institute.
FEATURES		Location/Qualifiers
Source		1..5564
gene		/organism="Homo sapiens"
CDS		/db_xref="taxon:9606"
		/clone="FLJ00152"
		/issue_type="spleen"
		/dev_stage="adult"
		/note="vector:PBuescriptII SK plus"
		1..5564
		/gene="FLJ00152"
		<1..4122
		/note="Start codon is not identified."
		/codon_start=1
		/evidence="not_experimental"
		/product="FLJ00152 protein"
		/protein_id="BAB84907.1"
		/db_xref="GI:18676510"
		/translation="GSDASNAMPVIFGKSSGPEFLOEYVATVYHNKSPDEYEYKIK LPKLTIVNHHLLFTPRHISCOOKACASVETLLGYNRPILMERLOTSCYCPVALLEK LPRNVSMSHAEVPLQNPPTKMAEGKVFNTIEVQAVSVYHNDHLEKFTPLCSLE SQTEPIRVLQDKISEMALEHLKLSICLNSNSVSVLFLHVLVDKFLQSLQVPMV IAGQTFANFQAFEEYVAIANSIHNKSKDLSKQHRNCLASVYHVRFLPEVQNDP KSCAPALDPPRSYHYTGRTSAAVSKSLQKROHNSNPDLAGTSHADAEVKIMS SKTADNCSRMYSYSGSSDAPSPAPAPAKKHIEHELALOMVYSGMVEYKY AMPEFLPLKSNAGOHYHNNDKSDSPRPFRPSPMDITTVYVNVSTELAAVLPQK ENBQAKRMNISLAFPLDLSLMDRGFVNLIRHCSQSLAKLSNLPILISRIEFLR ILSHBYHNLNLFENNADTAPSPSPSISONSSCSSFQDKIASMFDLSRQO HFTGGLFTEFLAALDAEBEGISVQKRAVSAIHLSHSDLPQCVAPENVKLTAL YLPVGLIIDLALPOLDFVADTRRYRTSGDEBEGAGAIQNVALAIAAGNNMLKT SGIYVSLSPKQVNMNADTPTNLMICFLIMKNADOSLRKIMADLSTQURLIDL LFTCVICFEYKKGSSDKSYTOYLQKSRVKAIRLEALLRGEAGHMRRAPENDIR FGLNENLKKKKQTHWRQANEKLDTRKLEDOEALISNLTFLGHAHLIILDMQNTIO ASSALDKDSLGGVLRVLVNSLNDQSTTYLTHCATLRALIAFLGDLFEEVBOC FDLGHVLIHHCSSMDVTRSOACATLYLMRFSFGATSNFARKQVYMSLVLGRA PDNEHHLRSLRTILAYSEEDTAMQMTPEFPQVEELCNLSIYDTVKMEFQEDP EMDIMLTRYAKSYOASPDRLTLQNMKEKTKKKCYEAMCVHAALVALEYLS LEPHASYLPGYSVSFQNISSNVLEESVSDTSLSPDGGCAGAGYFESGLVLLBOAA ELRSTGGLIETVYVERFKLPITLEAHRERKRLITLHRSKIQRAPDSIVNDRKMGCTY ERVFPFGSKFGLDDEDFVYKEPATIKLEPISIRNLAFLGGCFCAEFVYIDSDIPVD KTKLDPNKAVIQITFEVEPYFDEYEMKDRVYTEKKNLNRFMVTPETLEGPRDELH EGYRRNTVLTTHMAFPYIKIRISVIOKEEFVLPIEVAIEDMKKNPAVSSQH"
BASE COUNT		1533 a 1332 c 1314 g 1385 t
ORIGIN		
Query Match		14.3%: Score 71.4; DB 9; Length 5564;
Best Local Similarity		51.1%: Pred. No. 5.3e-11;
Matches 208; Conservative		0; Mismatches 181; Indels 18; Gaps 1;
Qy	2 GTTTACACATCCAAACCCAGATTTATGATGAGTTAAATAGAGTGGCCACT 61	
Db	82 GTTACATACCAATTAATAGTCTCTGACTTTATGAGAAGTGAATAATTAAGCTCCCGCT 141	
Qy	62 CAGCTGATGAAGACACACCTGTTCTCACATTTCTTCATGTCAGCTGCAACTCA 121	
Db	142 AAGCTCACAGTAATACACACCTCTGTCACCTTACCAATACACCTGTCAGCA---- 197	
Qy	122 AGTAAAGACGACGAGAGAGATGCTGTGAACCCAGTGGCTGCTCTGCTT 161	
Db	198 -----GAGCAAGAGAGCTCGGTGGAACCTCCTCGATATTCATGCGCTG 243	

TITLE Clasp-5 transmembrane protein
JOURNAL Patent: WO 0142296-A 1 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES Location/Qualifiers
source 1..7215
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="full length human CLASP-5 cDNA"
112..6159
CDS
/note="human CLASP-5"
/codon_start=1
/protein_id="CAC43733.1"
/db_xref="GI:14598201"
/translation="MTHLSNLSLDVLAQELGDFDTDDDDVYETPEKREKTLQPSLPEEY
ELDPHVRCVOTYIREMLIVNRKNSQSPETCGFKTSRDFDHTKLPKPFESLLEK
SEPAQOQPHNLNLCDSGKGYTACDFDRSLQPPKRLLENLQOYSADEPKQNE
ARKNRQAEFLALYPSVDEDAVEIRVPECPKHLNRLILVLLTKLELEPLA
SIALYDERKRISENDFGDLSDQFGFLRAHPVSAASQARSAAVESYTPSSDI
LVKLEKVLQOQDIDGCAEPTVTKESGSGKEIKLKQAESFCORLGYKMPA
MAPISLSEFNVSFLEREYVDVSVGSRSPGGERTLAOSRSLERALSLEINSGN
FKTSLSSSFQKQEGDLSDPELFLKPLADYKRSSLQRRYSKSLRLLEISNAPET
INCCITPEMLVKKPPEKNTIRPKHEILEPFRVYVPHYTRNLIIYYPQRLNFKVL
ASARNITIKIQMCGEDASNAAPVIFKSGSPREYLAQVYAVYHNKSPDYEEVATK
LPKLTIVNHLLEFTYHISCOQKQASVETLLGYSMTPLINERLQSGYCLPYALEK
LPNYSMSAKVPLONPPKMAEGHGVFNIEVOAVSVYTDHNLHEKFFTLCHSE
SOVPPRIVLDKISEMALHEHLKSTICLSSRLPLVLEFLHGLKLPOLSVQWV
IAGOTANSQPAESVAVAIANSLSKSDLSKDOGRCLIASVYHYVRLPEVORQVP
KSGAPTALDPRSHHTIGRTSAVAVSSLLDARWSSNPDLATGTHSADEEKNINS
SKIADNRCSRMYSYCSGSSDAISSPAARPAKSKHFEHLAQWVSTGVKMAQVH
HNMDKDSFRRTFRSDPMDDITIVNVVTSIALLVKPKQENKQAKNITSLAFVL
YDLISMDRGVFNLIIRYCSQKASLSNLTLSMLRFLICSHHYNLNLFPM
NADTAPSPCISISSONSSCSFODOKIASMPDTSRYROOHLITGLTFEILAAAD
AEGEISVORKAVSAIHSLSHDLDRCYKPKVKAIAVLPIGLTITDALPOLC
DFTVADRTRITSGSDEDEGAGAINONVALAIGNNPNLTSIVLSLPRKYNM
NADTTRNLMICFLWIMKNADQSLIRKVIADLPQNLNLDLITVLCETEKQOS
KROASTOVLOKSRDVKARLEALIRGEGRGMRRRAPGNDPFGLENLNMKKEQTH
WROANEKDKTKAELOEALISGNLTAHLIILDMQENIQAASSALDKXSLDGLYL
RVLVNSLNCDSOTYILTHCFATRLRLIKFGDLFEVEEQFDLCQHOVLHCSSMD
VTRSOACATYILIMRFSGATSNFAVRVMQVMTSLASIVGRAPDNENHRRSRTLL
AYSEEDTAMQMTPEPTQVEIELCNLSLITVYVMPREQDEPMLMLMTRIASYCA
SPDLRLWLNMAEKHNRKCYTEAMKLVLAALVAEYLSLMDHSLTPGVGSFON
ISSNVLESVSEDTLSPDEDEVCAGQFTESGLVGLAEQALLESSTGGLVETVNEY
KLIVPILAEHREFKLTLHSKLOAFDSIVNKHKRMFQYFRFGFSGFGLDQ
EEVYKEPATIKLEISHRLAEVGGCFGEVEVYKIDSTPDVKRDKPNKAYIQITFY
EYPRDEYKMDRVYFEKNPILRMETPTPLGRLHEDYRRNTVYTHHAP
YIKRTISYIOKEEFLPTPIEVALIDEMKTKTQLVAINOPEPDKMKLOVGSVQAT
VNQGLEAVOYFLAETPADPKYIRHNNKLRICEFELIRGCEAVENKRLITTAOREY
QOELKRNKNLKENLRLPMIERKIPILPYPIFRVBSQKRDSPHRSSFRCETQLSOGS"

BASE COUNT 2005 a 1734 c 1721 g 1755 t
ORIGIN

Query Match 14.3%; Score 71.4; DB 6; Length 7215;
Best Local Similarity 51.1%; Pred. No. 5,6e-11;

Matches 208; Conservative 0; Mismatches 181; Indels 18; Gaps 1;

QY 2 GTTTTACACCATCACCAGAAACCCAGAAATTTATGATGAGATTAAATAGTTGCCACT 61
DB 1759 GTTACATACCATATAAGTCTCTGCTTTTATGAAGAAGTGAATTAAGTCCCGCT 1818
QY 62 CAGCTGCATGAAGAAGCAGCACTGTTGCTCAGATTCCTTCATGCTGAGACAATCA 121
DB 1819 AAGCTCAGAGTAATACCAACCTCTGTCACCTTCTACCATATCAGCTGCAGCA---- 1874
QY 122 AGTAAAGGAAGCAGGAAGAGGATGCTGTTGAACCAAGTTGGCTACTGCTGCTT 181
DB 1875 -----GAGCAAGAGAGCCTCGTGGAACCTCTCGGAAATTCATTCATGCGCTG 1920
QY 182 CCCCTCTGAAGAAGAGGAGGTGTGACAGAGCAGACACATCCCGGCTCTCGGGAAC 241
DB 1921 CCAATTCTCTTAATGAACCTCTTCAAACTGGATCTTACTGCTCCCAAGTTGCCCTTGGAA 1980
QY 242 CTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGAGTGGGAGGAGCATATGATCGCGAA 301

DB 1981 AAATGGCACCACCACTACTCATGATCTTCTGCTGAGAAAGTCCCAATTCATCTCCC 2040
QY 302 ATTAATAGGCTAGTAGAGAGCAGCAGCAGCTGCTGAATTTTCCATGATCTGTTCTTACA 361
DB 2041 ATTAGTGGGCTAGAGCAGCAATTAAGGAGTATTAATATTGAATGACAGCTGTTCTTCT 2100
QY 362 GTGTATCTCAGATCAGCATGATTTACATTAATTTTTCAGACTGTCA 408
DB 2101 GTACACACCCAGAGCAACACACCTGAGAAAGTCTTCACTCCCTGCGCA 2147

RESULT 15
HSM803577
LOCUS HSM803577 9298 bp. mRNA linear PRI 12-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp66701117 (from clone DKFZp66701117).
ACCESSION AL832270
VERSION AL832270.1 GI:21732817
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 9298)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.
AUTHORS Direct Submission
TITLE Submitted (09-JUL-2002) 1, D-85764 Neuberger, GERMANY
JOURNAL Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
COMMENT Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by ACOMA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp66701117) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
source 1..9298
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp66701117"
/tissue_type="lymph node"
/clone_id="667 (synonym: hmo2). Vector psport1; host
DH10B; sites NotI + SalI
/dev_stage="adult"

polyA_signal 9259..9264
BASE COUNT 2546 a 2239 c 2161 g 2352 t
ORIGIN

Query Match 14.3%; Score 71.4; DB 9; Length 9298;
Best Local Similarity 51.1%; Pred. No. 5,9e-11;

Matches 208; Conservative 0; Mismatches 181; Indels 18; Gaps 1;

QY 2 GTTTTACACCATCACCAGAAACCCAGAAATTTATGATGAGATTAAATAGTTGCCACT 61
DB 3805 GTTACATACCATATAAGTCTCTGCTTTTATGAAGAAGTGAATTAAGTCCCGCT 3864
QY 62 CAGCTGCATGAAGAAGCAGCAGCAGCTGTTGCTCAGATTCCTTCATGCTGAGACAATCA 121
DB 3865 AAGCTCAGAGTAATACCAACCTCTGTCACCTTCTACCATATCAGCTGTGACAGCA---- 3920
QY 122 AGTAAAGGAAGCAGGAAGAGGATGCTGTTGAACCAAGTTGGCTACTGCTGCTT 181
DB 3921 -----GAGCAAGAGAGCCTCGTGGAACCTCTCGGAAATTCATTCATGCGCTG 3966
QY 182 CCCCTCTGAAGAAGAGGAGGTGTGACAGAGCAGACACATCCCGGCTCTCGGGAAC 241
DB 3967 CCAATTCTCTTAATGAACGCTTCAAACTGGATCTTACTGCTCCCAAGTTGCCCTTGGAA 4026
QY 242 CTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGAGTGGGAGGAGCATATGATCGCGAA 301
DB 4027 AAATTGCCACCACTACTCATGATCTTCTGCTGAGAAAGTCCCAATTCATGATCTGCC 4086

```

OY 302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAATTTCCACTGATCTGGTTTCTACA 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4087 ATTAAAGTGGGCTGAAGGACATTAAGGAGTATTTAATATTGAAGTGCACAGCTGTTCTCT 4146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 362 GTGATATCTCAGGATCAGCATTTACATTAATTTTTCAGTACTGTCA 408
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4147 GTACACACCCAGGACACCACTGTGAGAAAGTTCTTCACCCCTCTGCCA 4193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: February 7, 2003, 08:03:14
 Job time : 1332.67 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:36:04 ; Search time 127.83 seconds
(without alignments)
8808.598 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agtttacacatcaccaaa.....agccacgtgatgcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	4807	21	AAC87972 Human CLASP-2 nucl
2	500	100.0	4807	21	AAC87973 Human CLASP-2A nuc
3	500	100.0	4807	24	ABR84966 DNA encoding cadhe
4	500	100.0	4807	24	ABR84973 DNA encoding cadhe
5	499	99.8	4806	24	ABR84964 DNA encoding cadhe
6	499	99.8	5048	24	ABR84965 DNA encoding cadhe
7	498.4	99.7	5862	24	ABR84970 DNA encoding cadhe
8	498.4	99.7	6816	21	AAC74524 Human ORF79 p
9	498.4	99.7	7506	22	AA019118 Angiogenesis assoc

10	496.8	99.4	6791	24	ABR85003 DNA encoding cadhe
11	486.4	97.3	4898	21	AAC87974 Preliminary CLASP-
12	486.4	97.3	4898	21	AAC87975 Preliminary CLASP-
13	489.4	97.3	4898	21	AAC87976 Preliminary CLASP-
14	486.4	97.3	4898	21	AAC87977 Preliminary CLASP-
15	486.4	97.3	4898	21	AAC87978 Preliminary CLASP-
16	486.4	97.3	4898	21	AAC87979 Preliminary CLASP-
17	486.4	97.3	4898	21	AAC87980 Preliminary CLASP-
18	486.4	97.3	4898	21	AAC87981 Preliminary CLASP-
19	486.4	97.3	4898	24	ABR84992 DNA encoding cadhe
20	486.4	97.3	4898	24	ABR84993 DNA encoding cadhe
21	486.4	97.3	4898	24	ABR84994 DNA encoding cadhe
22	486.4	97.3	4898	24	ABR84995 DNA encoding cadhe
23	486.4	97.3	4898	24	ABR84996 DNA encoding cadhe
24	486.4	97.3	4898	24	ABR84997 DNA encoding cadhe
25	486.4	97.3	4898	24	ABR84998 DNA encoding cadhe
26	486.4	97.3	4898	24	ABR84999 DNA encoding cadhe
27	204.4	40.9	6454	22	AA508334 Human CDNA encodin
28	173.8	34.8	5688	21	AAC87969 Human CLASP relate
29	173.8	34.8	5688	21	AA414825 Human CDNA encodin
30	173.8	34.8	7277	24	AA518951 Human CDNA encodin
31	117.2	23.4	5214	21	AAC87968 Mouse CLASP-1 nucl
32	117.2	23.4	5214	21	AA414824 DNA encoding a mur
33	85	17.0	4393	22	AA508357 Human CDNA encodin
34	71.4	14.3	7215	22	AA507373 Human CDNA encodin
35	65.8	14.0	665	24	ABR53548 Human eosinophil-m
36	54.2	10.8	5955	23	AB111149 Drosophila melanog
37	46	9.2	6372	22	AAH43851 Human CLASP-7 enco
38	42.4	8.5	5589	23	AB117737 Drosophila melanog
39	42.4	8.5	9389	23	AB117736 Drosophila melanog
40	41	8.2	6828	22	AAH41934 Human CLASP-3 cDNA
41	40.4	8.1	211	22	AAH43868 Human CLASP-7 intr
42	40.4	8.1	3023	22	AAH17551 Human CDNA sequenc
43	38.6	7.7	9842	23	AB111148 Enterococcus faeca
44	34.6	6.9	6691	20	AAI33025 Eosinophil activat
45	34	6.8	1851	21	AA52623

ALIGNMENTS

RESULT 1					
ID	AAC87972	standard;	CDNA:	4807	BP.
AC	AAC87972;				
DT	07-MAR-2001	(first entry)			
DE	Human CLASP-2 nucleotide sequence.				
KW	CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;				
KW	cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;				
KW	immunomodulatory; antiinflammatory; antitumor; cytostatic;				
KW	hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;				
KW	hypersensitivity; transplantation rejection response; immunodeficiency;				
KW	proliferation; differentiation; inflammatory response; arthritis;				
KW	inflammatory bowel disease; hematopoietic cell; blood protein disorder;				
KW	anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;				
KW	endometriosis; pregnancy induced hypertension; ss.				
OS	Homo sapiens.				
XX					
XX					
PN	WO200061747-A2.				
PD	19-OCT-2000.				
XX					
PF	13-APR-2000;	2000WO-US10158.			
XX					
PR	14-APR-1999;	99US-0129171.			
XX					
PR	14-MAY-1999;	99US-0134114.			
XX					
PR	14-MAY-1999;	99US-0134117.			
XX					
PR	14-MAY-1999;	99US-0134118.			

```
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
XX
PI Lu PS:
XX
XX WPI; 2000-619230/59.
DR P-PSDB; AAB36527.
XX
PT Isolated cadherin-1-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
XX Example 1; Fig 1; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antiarthritic, cytostatic, hypotensive, antineumatic,
XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX immune response in a subject may also be inhibited by administering an
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX proteins and antibodies can be used to prevent or treat a CLASP-2
XX mediated disease, such as an autoimmune disease caused or exacerbated
XX by increased activity of TH1 cells. They can also be used to treat
XX hypersensitivities, prevent transplantation rejection responses and
XX augment immune responsiveness in immunodeficiency states, inhibit
XX proliferation and differentiation of cells involved in an inflammatory
XX response e.g. arthritis, inflammatory bowel disease and increase
XX differentiation and proliferation of haematopoietic cells e.g. to treat
XX anaemia, thrombocytopaenia and other blood protein disorders. Disorders
XX created by disrupting CLASP-2 function include multiple sclerosis,
XX rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
XX The present sequence encodes human CLASP-2, which is used in the
XX exemplification of the present invention.
XX
SO Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other:
XX
Query Match 100.0%; Score 500; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.1e-157;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGTTTACACCATCAGCAAAACCCAGAAATTTATGATGATTAAATGAGTTGCCAC 60
DB 1 AGTTTACACCATCAGCAAAACCCAGAAATTTATGATGATTAAATGAGTTGCCAC 60
XX
QY 61 TCAGCTGATGAAAGCAGCACCCTGTTGCTCATCTTCCATGTCAGTGGACACATC 120
DB 61 TCAGCTGATGAAAGCAGCACCCTGTTGCTCATCTTCCATGTCAGTGGACACATC 120
XX
QY 121 AAGTAAAGACGACGACGAGAGAGGATGTCGTTGAACCCAGTGGCTACTCTGCT 180
DB 121 AAGTAAAGACGACGACGAGAGAGGATGTCGTTGAACCCAGTGGCTACTCTGCT 180
XX
QY 181 TCCCTCTCTGAAAGAGCGAGGCTGTGACAAAGCAGACGACATCCGGTCTGGCGAA 240
DB 181 TCCCTCTCTGAAAGAGCGAGGCTGTGACAAAGCAGACGACATCCGGTCTGGCGAA 240
XX
QY 241 CCTTCTTGGGGGTATCTGGCTACCAAGAGTGGAGTGGCAGGACATATGATCGCGA 300
DB 241 CCTTCTTGGGGGTATCTGGCTACCAAGAGTGGAGTGGCAGGACATATGATCGCGA 300
XX
QY 301 AATTAATGGGTAGATGAGGACAGCACTGCTGAAAATTTCCATCTCATCTGGTTTAC 360
DB 301 AATTAATGGGTAGATGAGGACAGCACTGCTGAAAATTTCCATCTCATCTGGTTTAC 360
XX
QY 361 AGTGTACTCTGAGATCAGCATTTTACATATATTTTTCAGTACTGCGAANAACGATC 420
DB 361 AGTGTACTCTGAGATCAGCATTTTACATATATTTTTCAGTACTGCGAANAACGATC 420
XX
```

```
DB 361 AGTGTACTCTGAGATCAGCATTTTACATATATTTTTCAGTACTGTCGAAAACCGATC 420
QY 421 TGGAGCCCAAGCCTTAGGAAAAGAACTTGTAAAGTACCTTAAGAGTGTGATGCCATGA 480
DB 421 TGGAGCCCAAGCCTTAGGAAAAGAACTTGTAAAGTACCTTAAGAGTGTGATGCCATGA 480
QY 481 AGGCCACGTGATGATCGCCT 500
DB 481 AGGCCACGTGATGATCGCCT 500
XX
RESULT 2
AAC87973
ID AAC87973 standard; cDNA; 4807 BP.
XX
AC AAC87973;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CLASP-2A nucleotide sequence.
XX
XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
XX cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
XX immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
XX hypotensive; antineumatic; antianaemic; haemostatic; neuroprotective;
XX hypersensitivity; transplantation rejection response; immunodeficiency;
XX proliferation; differentiation; inflammatory response; arthritis;
XX inflammatory bowel disease; haematopoietic cell; blood protein disorder;
XX anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
XX endometriosis; pregnancy induced hypertension; ss.
XX
XX Homo sapiens.
XX
XX W0200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-Apr-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
XX
PI Lu PS:
XX
XX WPI; 2000-619230/59.
DR P-PSDB; AAB36528.
XX
PT Isolated cadherin-1-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
XX Example 1; Fig 2B; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antiarthritic, cytostatic, hypotensive, antineumatic,
XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX immune response in a subject may also be inhibited by administering an
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX proteins and antibodies can be used to prevent or treat a CLASP-2
XX mediated disease, such as an autoimmune disease caused or exacerbated
XX by increased activity of TH1 cells. They can also be used to treat
```



```
QY 1 AGTTTACACCATCACCACCAACCCAGAAATTTATGATGAGATTAAATAGAGTTGCCAC 60
DB 1 AGTTTTCACCATCACCACCAACCCAGAAATTTATGATGAGATTAAATAGAGTTGCCAC 60
QY 61 TCAGCTGCATGAAAGACCAACCTGTTGCTTCACATTCCTCCATGTCAGCTGTGACAATC 120
DB 61 TCAGCTGCATGAAAGACCAACCTGTTGCTTCACATTCCTCCATGTCAGCTGTGACAATC 120
QY 121 AAGTAAAGGAAGCAGCAAGAGAGGATGCTGTGAACCCAGAGTTGGCTACTCTGGCT 180
DB 121 AAGTAAAGGAAGCAGCAAGAGAGGATGCTGTGAACCCAGAGTTGGCTACTCTGGCT 180
QY 181 TCCCTCTCTTAAGACGAGAGGTGTGACACGACGACACATCCCGTCTCGCGGAA 240
DB 181 TCCCTCTCTTAAGACGAGAGGTGTGACACGACGACACATCCCGTCTCGCGGAA 240
QY 241 CCTTCTCTGGGGCTATCTTGGCTACCAAGAGCTTGGGATGGCAGCATTTATGTCCGA 300
DB 241 CCTTCTCTGGGGCTATCTTGGCTACCAAGAGCTTGGGATGGCAGCATTTATGTCCGA 300
QY 301 AATTAAATGGGTAGTAGGAGGCAAGCCACTGTAATAATTTCCACTCATCTGGTTCTAC 360
DB 301 AATTAAATGGGTAGTAGGAGGCAAGCCACTGTAATAATTTCCACTCATCTGGTTCTAC 360
QY 361 AGTGTATACTCAGAGATCAGATTACATATTTTTCAGTACTGTCAAGAAACCGAATC 420
DB 361 AGTGTATACTCAGAGATCAGATTACATATTTTTCAGTACTGTCAAGAAACCGAATC 420
QY 421 TGGAGCCCAAGCCTTAGGAAGCACTTGTAAAGTACCTTAGAGTCTGCATGCGAWGA 480
DB 421 TGGAGCCCAAGCCTTAGGAAGCACTTGTAAAGTACCTTAGAGTCTGCATGCGAWGA 480
QY 481 AGGCGACGTGATGTCGCCT 500
DB 481 AGGCGACGTGATGTCGCCT 500

RESULT 4
ABK84973
ID ABK84973 standard; DNA: 4807 BP.
XX
AC ABK84973;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform.
XX
KW Human: autoimmune disease; haematopoietic disorder; Digesorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARW/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
```

```
DR P-PSDB; ABG61686.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis
XX
PS Example 4; Figure 6A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digesorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences
CC and PCR primers of the invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other:
Query Match 100.0%; Score 500; DB 24; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3; 1e-157;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTTTACACCATCACCACCAACCCAGAAATTTATGATGAGATTAAATAGAGTTGCCAC 60
DB 1 AGTTTTCACCATCACCACCAACCCAGAAATTTATGATGAGATTAAATAGAGTTGCCAC 60
QY 61 TCAGCTGCATGAAAGACCAACCTGTTGCTTCACATTCCTCCATGTCAGCTGTGACAATC 120
DB 61 TCAGCTGCATGAAAGACCAACCTGTTGCTTCACATTCCTCCATGTCAGCTGTGACAATC 120
QY 121 AAGTAAAGGAAGCAGCAAGAGAGGATGCTGTGAACCCAGAGTTGGCTACTCTGGCT 180
DB 121 AAGTAAAGGAAGCAGCAAGAGAGGATGCTGTGAACCCAGAGTTGGCTACTCTGGCT 180
QY 181 TCCCTCTCTTAAGACGAGAGGTGTGACACGACGACACATCCCGTCTCGCGGAA 240
DB 181 TCCCTCTCTTAAGACGAGAGGTGTGACACGACGACACATCCCGTCTCGCGGAA 240
QY 241 CCTTCTCTGGGGCTATCTTGGCTACCAAGAGCTTGGGATGGCAGCATTTATGTCCGA 300
```

```

Db      | 241 CCTCTCTGGGCTATCTTGGCTACCAAGACCTTGGGATGGGCGGCGATATGTTGCCGGA 300
Oy      | 301 AATTAATAGGTAGATGAGGCAAGCCACCTGCTGAAATTTCCACTCATCTGTTTCTAC 360
Db      | 301 AATTAATAGGTAGATGAGGCAAGCCACCTGCTGAAATTTCCACTCATCTGTTTCTAC 360
Oy      | 361 AGTATATCTAGATGATGCACTTTACATTAATTTTTTCCAGTACTGTGAGAAACCGAATC 420
Db      | 361 AGTATATCTAGATGATGCACTTTACATTAATTTTTTCCAGTACTGTGAGAAACCGAATC 420
Oy      | 421 TGGAGCCCAAGCCTTAGAAGCAAGCTTGAAGTACCTTAAGAGTCTGATGCGATGGA 480
Db      | 421 TGGAGCCCAAGCCTTAGAAGCAAGCTTGAAGTACCTTAAGAGTCTGATGCGATGGA 480
Oy      | 481 AGGCCACGTGATGATGCGCT 500
Db      | 481 AGGCCACGTGATGATGCGCT 500

```

RESULT 5 ABK84964

ID ABR84964 standard; cDNA; 4806 BP.

AC ABR84964;

DT 13-Aug-2002 (first entry)

DE DNA encoding cadherin-like asymmetry protein (CLASP).

Human: autoimmune disease; hematopoietic disorder; Diegoorge syndrome;
blood protein disorder; agammaglobulinemia; dysgammaaglobulinemia;
ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
Kw hemocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
hemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
rheumatoid arthritis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
autoimmune pulmonary inflammation; organ rejection; inflammation;
CLASP; gene; ss.

OS Homo sapiens.

PN WO20023117-A2.

PD 18-Apr-2002.

PF 15-Oct-2001; 2001WO-US32202.

PR 13-Oct-2000; 2000US-0687837.

PA (ARBO-) ARBOR VITA CORP.

PA (GARMA/) GARMAN J D.

PA (CAND/) CANDIA A F.

PI Lu PS;

DR WPI: 2002-416861/44.

DR P-PSDB; ABG61670.

PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
an immune response, and for treating multiple sclerosis, rheumatoid
arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
and sepsis

PS Disclosure; Figure 1; 245bp; English.

XX The invention relates to an isolated polypeptide (I) comprising an amino
XX acid sequence that has 90 % sequence identity to one of the human
XX CLASP-2 polypeptides (e.g., antisense or
XX ribozyme-mediated inhibition), for gene knock-out, etc. The CLASP-2
XX antibodies or are used as therapeutic polypeptides. The CLASP-2
XX polynucleotide or fragments can be used in diagnostics (e.g., as probes
XX for CLASP-2 expression), as a lymphocyte marker and for therapeutic
XX purposes. CLASP-2 polynucleotides can construct transgenic and knock-out
XX animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
XX polynucleotides can screen for CLASP-2 agonists and antagonists.
XX CLASP-2 polypeptides or polynucleotides can treat deficiencies or
XX disorders of the immune system, by activating or inhibiting the
XX activation, differentiation of immune cells and can treat or detect
XX deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
XX or polynucleotides can increase differentiation and proliferation of
XX haematopoietic cells, including the pluripotent stem cells to treat those
XX disorders associated with a decrease in certain (or many) types of
XX haematopoietic cells (e.g., immunologic deficiency syndromes including
XX dysgammaaglobulinemia, ataxia telangiectasia, common variable
XX immunodeficiency, Diegoorge syndrome, lymphopenia, thrombocytopenia, or
XX haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
XX detect autoimmune diseases, e.g., Addison's disease, hemolytic anaemia,
XX Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
XX endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
XX inflammation. CLASP-2 can be used to treat anaphylaxis or
XX hypersensitivity to an antigenic molecules, organ rejection or graft-
XX versus-host disease (GVHD) and inflammation. ABR84922-ABR85018
XX represent cadherin-like asymmetry protein (CLASP) coding sequences and
XX PCR primers of the invention.

or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
the autoimmune disease is caused or exacerbated by increased activity
of Th1 cells. CLASP-2 polynucleotides are useful as probes or primers for
detection or inhibition of CLASP-2 expression (e.g., antisense or
ribozyme-mediated inhibition), for gene knock-out, etc. The CLASP-2
antibodies or are used as therapeutic polypeptides, produce anti-CLASP-
polynucleotide or fragments can be used in diagnostics (e.g., as probes
for CLASP-2 expression), as a lymphocyte marker and for therapeutic
purposes. CLASP-2 polynucleotides can construct transgenic and knock-out
animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
polynucleotides can screen for CLASP-2 agonists and antagonists.
CLASP-2 polypeptides or polynucleotides can treat deficiencies or
disorders of the immune system, by activating or inhibiting the
activation, differentiation of immune cells and can treat or detect
deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
or polynucleotides can increase differentiation and proliferation of
haematopoietic cells, including the pluripotent stem cells to treat those
disorders associated with a decrease in certain (or many) types of
haematopoietic cells (e.g., immunologic deficiency syndromes including
dysgammaaglobulinemia, ataxia telangiectasia, common variable
immunodeficiency, Diegoorge syndrome, lymphopenia, thrombocytopenia, or
haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
detect autoimmune diseases, e.g., Addison's disease, hemolytic anaemia,
Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
inflammation. CLASP-2 can be used to treat anaphylaxis or
hypersensitivity to an antigenic molecules, organ rejection or graft-
versus-host disease (GVHD) and inflammation. ABR84922-ABR85018
represent cadherin-like asymmetry protein (CLASP) coding sequences and
PCR primers of the invention.

Sequence 4806 BP; 1352 A; 1117 C; 1136 G; 1201 T; 0 other:

Query Match 99.8%; Score 499; DB 24; Length 4806;
Best Local Similarity 100.0%; Pred. No. 6,7e-157;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      | 2 GTTTACACCATCCCAAAACCCAGATTTTATGATGATTAATTAATGAGTGGCCACT 61
Db      | 1 GTTTACACCATCCCAAAACCCAGATTTTATGATGATTAATTAATGAGTGGCCACT 60
Oy      | 62 CAGCTGCATGAAGAAGCACCCTGTTGCTCACATCTTCATGCTGAGTGCACATCA 121
Db      | 61 CAGCTGCATGAAGAAGCACCCTGTTGCTCACATCTTCATGCTGAGTGCACATCA 120
Oy      | 122 AGTAAAGAACGACGAAGAGGATGCTTGAACCCAAAGTTGGCTACTCTGGCTT 181
Db      | 121 AGTAAAGAACGACGAAGAGGATGCTTGAACCCAAAGTTGGCTACTCTGGCTT 180
Oy      | 182 CCCCTCCTGAAGACGGAAGGGTGTGACCAACGAGCAACATCCCGCTCGGCGAAC 241
Db      | 181 CCCCTCCTGAAGACGGAAGGGTGTGACCAACGAGCAACATCCCGCTCGGCGAAC 240
Oy      | 242 CTTCTTGGGCTATCTTGGTACCAAGAGCTTGGATGGGAGGACATTAATGATGTC 301
Db      | 241 CTTCTTGGGCTATCTTGGTACCAAGAGCTTGGATGGGAGGACATTAATGATGTC 300
Oy      | 302 ATTAATAGGTAGATGAGGCAAGCCACTGCTGAATTTTCCATCATCTGTTCTACA 361
Db      | 301 ATTAATAGGTAGATGAGGCAAGCCACTGCTGAATTTTCCATCATCTGTTCTACA 360
Oy      | 362 GTGTATACAGATCAGATTAATTAATTTTCCAGTCTGCGAAGAAACCGAATCT 421
Db      | 361 GTGTATACAGATCAGATTAATTAATTTTCCAGTCTGCGAAGAAACCGAATCT 420
Oy      | 422 GGAGCCCAAGCCTTAGAAGCAACTTGAAGTACCTTAAGAGTCTGATGATGAA 481
Db      | 421 GGAGCCCAAGCCTTAGAAGCAACTTGAAGTACCTTAAGAGTCTGATGATGAA 480
Oy      | 482 GGCCACGTGATGATGCGCT 500

```

DB 481 GGCACGCTATGATCCCT 499

RESULT 6
ABK84965
ABK84965 standard; cDNA: 5048 BP.

AC ABK84965;
DT 13-AUG-2002 (first entry)
XX DNA encoding cadherin-like asymmetry protein (CLASP) isoform.

XX Human; autoimmune disease; haematopoietic disorder; Digorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.

XX Homo sapiens.
OS
XX MO200231117-A2.
PN
XX 18-APR-2002.
PD
XX 15-OCT-2001; 2001MO-US32202.
PF
XX 13-OCT-2000; 2000US-0687837.
PR
XX
XX (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
PI
XX Lu PS;
XX WPI; 2002-416861/44.
DR P-PSDB; ABG61871.
DR
XX
XX New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis

PS Disclosure: Figure 2; 245pp; English.

XX The invention relates to an isolated polypeptide (I) comprising an amino
XX acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TMI cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of

CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, hemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.

XX
XX Sequence 5048 BP: 1403 A; 1187 C; 1215 G; 1243 T; 0 other;
SQ

Query Match 99.8%; Score 499; DB 24; Length 5048;
Best local Similarity 100.0%; Pred. No. 6.9e-157;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTTTACACATCACCAAAACCCAGAAATTTATGATGAGATTAAATAGATGCCACT 61
DB 1 GTTTACACATCACCAAAACCCAGAAATTTATGATGAGATTAAATAGATGCCACT 60

QY 62 CAGCTGCATGAAAGACACACCTGTGCTCATCTTCCATGTCAGCTGACCAACTCA 121
DB 61 CAGCTGCATGAAAGACACACCTGTGCTCATCTTCCATGTCAGCTGACCAACTCA 120

QY 122 AGTAAAGAACACGAGAAAGAGGATGTCGTTGAAACCCAGTTGGCTACTCGTGCCT 181
DB 121 AGTAAAGAACACGAGAAAGAGGATGTCGTTGAAACCCAGTTGGCTACTCGTGCCT 180

QY 182 CCCCTCTGAAAGACGAGGAGGTGTGACAGCAGCAGCATCCGGTCTGGCGAAC 241
DB 181 CCCCTCTGAAAGACGAGGAGGTGTGACAGCAGCAGCATCCGGTCTGGCGAAC 240

QY 242 CTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGATGGGACGACATTTATGGCCGAA 301
DB 241 CTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGATGGGACGACATTTATGGCCGAA 300

QY 302 ATTAATGGTAGATGAGGACGACCACTGCTGAAAATTTCCACTCATCTGGTTTACA 361
DB 301 ATTAATGGTAGATGAGGACGACCACTGCTGAAAATTTCCACTCATCTGGTTTACA 360

QY 362 GTTATACTCAGATCAGCATTTACATATTTTTTCCAGTACTGTGCAAAACCGAATCT 421
DB 361 GTTATACTCAGATCAGCATTTACATATTTTTTCCAGTACTGTGCAAAACCGAATCT 420

QY 422 GSAGCCCAAGCCTTAGAAGCAACTGTAAAGTACCTTAGAGTGTGATGCGATGGAA 481
DB 421 GSAGCCCAAGCCTTAGAAGCAACTGTAAAGTACCTTAGAGTGTGATGCGATGGAA 480

QY 482 GGCACGATGATGCGCT 500
DB 481 GGCACGATGATGCGCT 499

RESULT 7
ABK84970
ABK84970 standard; cDNA: 5862 BP.

AC ABK84970;
DT 13-AUG-2002 (first entry)
XX DNA encoding cadherin-like asymmetry protein (CLASP) isoform #5.

XX Human; autoimmune disease; haematopoietic disorder; Digorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;

KM haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
 KM endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
 KM autoimmune pulmonary inflammation; organ rejection; inflammation;
 KM CLASP; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200231117-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-US32202.
 XX
 PR 13-OCT-2000; 2000US-0687837.
 XX
 PA (ARBO-) ARBOR VITAE CORP.
 PA (GARMA) GARMAN J D.
 PA (CAND/) CANDIA A F.
 PI
 PI Lu PS:
 XX
 DR WPI: 2002-416861/44.
 DR P-PSDB: AB661676.
 XX
 PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
 PT an immune response, and for treating multiple sclerosis, rheumatoid
 PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
 PT and sepsis
 PS Disclosure: Figure 3A: 245pp; English.
 XX
 CC The invention relates to an isolated polypeptide (I) comprising an amino
 CC acid sequence that has 90 % sequence identity to one of the human
 CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
 CC sequences (PS). (I) is useful for identifying a compound or agent that
 CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
 CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
 CC inhibiting an immune response in a subject. A pharmaceutical composition
 CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
 CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
 CC the autoimmune disease is caused or exacerbated by increased activity
 CC of Th1 cells. CLASP-2 polynucleotides are useful as probes or primers for
 CC detection or inhibition of CLASP-2 expression (e.g., antisense or
 CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
 CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
 CC antibodies or are used as therapeutic polypeptides. The CLASP-2
 CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
 CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
 CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
 CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
 CC polynucleotides can screen for CLASP-2 agonists and antagonists.
 CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
 CC disorders of the immune system, by activating or inhibiting the
 CC activation, differentiation of immune cells and can treat or detect
 CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
 CC or polynucleotides can increase differentiation and proliferation of
 CC haematopoietic cells, including the pluripotent stem cells to treat those
 CC disorders associated with a decrease in certain (or many) types of
 CC haematopoietic cells e.g., immunologic deficiency syndromes including
 CC blood protein disorders (e.g., agammaglobulinemia, common variable
 CC dysgammaglobulinemia, ataxia telangiectasia, immunodeficiency, DiGeorge syndrome, lymphopenia, thrombocytopenia, or
 CC immunodeficiency. CLASP-2 polynucleotides or polypeptides can treat or
 CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
 CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
 CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
 CC inflammation. CLASP-2 can be used to treat anaphylaxis or
 CC hypersensitivity to an antigenic molecules, organ rejection or graft-
 CC versus-host disease (GVHD) and inflammation. ABX84922-ABX85018
 CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
 CC PCR primers of the invention.
 XX
 SO Sequence 5862 BP; 1670 A; 1327 C; 1341 G; 1524 T; 0 other;

Query Match 99.7%; Score 498.4; DB 24; Length 5862;
 Best Local Similarity 99.8%; Pred. No. 1.2e-156;
 Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AGTTTACACCATCACCACCAAGAAATTTATGATGATTAATATAGATTGCCAC 60
 DB 495 AGTTTACACCATCACCACCAAGAAATTTATGATGATTAATATAGATTGCCAC 554
 OY 61 TCAGCTGCATGAAAGGACCCACCTGTCATCATCTTCATGTCAGCTGTCACACTC 120
 DB 555 TCAGCTGCATGAAAGGACCCACCTGTCATCATCTTCATGTCAGCTGTCACACTC 614
 OY 121 AAGTAAAGGACGACGAAAGAGGATGCTGTCGAAACCCAGTTGGCTACTCGCT 180
 DB 615 AAGTAAAGGACGACGAAAGAGGATGCTGTCGAAACCCAGTTGGCTACTCGCT 674
 OY 181 TCCTCTCTGAAAGGACGAAAGGATGTCAGCAAGCAGCAGCATCCCGTCTGGCGAA 240
 DB 675 TCCTCTCTGAAAGGACGAAAGGATGTCAGCAAGCAGCAGCATCCCGTCTGGCGAA 734
 OY 241 CCTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGATGGGACGCAATATGTCGCGA 300
 DB 735 CCTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGATGGGACGCAATATGTCGCGA 794
 OY 301 AATTAAATGGGTATGATGAGGCAAGCAGCAGCTGTGAAATTTCCACTATCTGTTTCTAC 360
 DB 795 AATTAAATGGGTATGATGAGGCAAGCAGCAGCTGTGAAATTTCCACTATCTGTTTCTAC 854
 OY 361 AGTGTATCTACAGATCAGATCATTTACATTAATTTTCCAGTACGTGCAAGAACCGAATC 420
 DB 855 AGTGTATCTACAGATCAGATCATTTACATTAATTTTCCAGTACGTGCAAGAACCGAATC 914
 OY 421 TGGAGGCCAAGCCTTATGAAAGCAAGCACTGTAAAGTCTTAAGATCTGATGCGATGGA 480
 DB 915 TGGAGGCCAAGCCTTATGAAAGCAAGCACTGTAAAGTCTTAAGATCTGATGCGATGGA 974
 OY 481 AGGCGACGATGATGCGCT 500
 DB 975 AGGCGACGATGATGCGCT 994
 RESULT 8
 AAC74524
 ID AAC74524 standard; cDNA; 6816 BP.
 XX
 AC AAC74524;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORF79 polynucleotide sequence SEQ ID NO:157.
 XX
 XX Human: open reading frame; ORF; detection; cytosolic; hepatotropic;
 KM vulnary; antiproliferative; antiparkinsonian; neuroprotective;
 KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KM antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2000058473-A2.
 PD 05-OCT-2000.

```
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI: 2000-602362/57.
DR P-PSDB; AAB40315.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 502-506; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiproliferic; antiparinsonian; nootropic; neuroprotective;
CC osteoplastic; anticonvulsant; antichronic; immunosuppressive;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antihydroid; and antinaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX
SO Sequence 6816 BP: 1977 A; 1523 C; 1557 G; 1757 T; 2 other;
Query Match 99.7%; Score 498.4; DB 21; Length 6816;
Best Local Similarity 99.8%; Pred. No. 1.3e-156;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGTTTACACCATCACCAAAACCCAGAAATTTATGATGATTAAATAGAGTCCGCAC 60
DB 1443 AGTTTACACCATCACCAAAACCCAGAAATTTATGATGATTAAATAGAGTCCGCAC 1502
QY 61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTTCCATGTCAGCTGACAACTC 120
DB 1503 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTTCCATGTCAGCTGACAACTC 1562
QY 121 AACTTAAGGAAGACGAAGAAGAGGATGCTGTGAACCCAGTGGTACTCCTGGCT 180
DB 1563 AACTTAAGGAAGACGAAGAAGAGGATGCTGTGAACCCAGTGGTACTCCTGGCT 1622
QY 181 TCCCTCTCTGAAGACGAAGAGGTGTGACAAAGCAGACATCCCGTCTCGCGCAA 240
DB 1623 TCCCTCTCTGAAGACGAAGAGGTGTGACAAAGCAGACATCCCGTCTCGCGCAA 1682
QY 241 CCTTCCTTGGGGCTATCTTGCTTACCAAGAGCTTGGGATGGGAGCATTAAGTTCGGA 300
DB 1683 CCTTCCTTGGGGCTATCTTGCTTACCAAGAGCTTGGGATGGGAGCATTAAGTTCGGA 1742
QY 301 AATTAATGGGTGATGAGAGCAAGCACTGTGAATAATTCACATCACTGGTTCTAC 360
DB 1743 AATTAATGGGTGATGAGAGCAAGCACTGTGAATAATTCACATCACTGGTTCTAC 1802
QY 361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTCCAGTACTGTGAGAAAACCGAATC 420
```

```
DB 1803 AGTGTATACTCAGGATCAGCATTTTACATTAATTTTCCAGTACTGTGAGAAAACCGAATC 1862
QY 421 TGGAGCCCAAGCCTTAGAAGCAAGCTGTGAAGTAAAGTCTGATGCATGGA 480
DB 1863 TGGAGCCCAAGCCTTAGAAGCAAGCTGTGAAGTAAAGTCTGATGCATGGA 1922
QY 481 AGGCCACGTGATGATCGCCT 500
DB 1923 AGGCCACGTGATGATCGCCT 1942
RESULT 9
AAB19118
ID AAB19118 standard; cDNA: 7506 BP.
AC AAB19118;
XX
XX 18-DEC-2001 (first entry)
DT
XX
DE Angiogenesis associated human thyroid regulated transcript (hTRG) cDNA.
XX
XX Angiogenesis associated protein; AAP; cytostatic; cardiant; gene therapy;
XX ophthalmological; vulnery; myocardial infarction; macular degeneration;
XX diabetic retinopathy; angiogenesis; wound healing; prophylactic; vaccine;
XX rheumatoid arthritis; psoriasis; drug screening; tumour; transplantation;
XX cancer; therapeutic; diagnostic; human; thyroid regulated transcript;
XX TRG; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 43..6366
XX FT /*tag= a
XX FT /product= "human thyroid regulated transcript (hTRG)"
XX
XX WO200170808-A2.
XX
XX 27-SEP-2001.
XX
XX 22-MAR-2001; 2001WO-US09609.
XX
XX 22-MAR-2000; 2000US-191134P.
XX
XX (CURA-) CURAGEN CORP.
XX PA (GETH ) GENENTECH INC.
XX
XX Rastelli LK, Gerritsen M;
XX
XX WPI: 2001-602775/68.
XX
XX P-PSDB; AAE11889.
XX
XX Novel angiogenesis associated polypeptides and polynucleotides encoding
XX the polypeptides, useful for modulating angiogenesis and for treating
XX tumours and cancers -
XX
XX Claim 6; Page 18-21; 159pp; English.
XX
XX The invention relates to angiogenesis associated proteins (AAP) and their
XX corresponding cDNA molecules, which are useful for modulating
XX angiogenesis. AAP proteins and nucleic acids are useful for promoting
XX wound healing, for example after organ transplantation and in the
XX treatment of tumours, myocardial infarction, cancers, diabetic
XX retinopathy, macular degeneration, psoriasis and rheumatoid arthritis.
XX AAP proteins and DNA's are useful in potential prophylactic and
XX therapeutic applications implicated in a variety of disorders including
XX those related to angiogenesis, and also in diagnostic applications.
XX AAP cDNA is also useful in gene therapy. The invention also relates to
XX a method for screening a tissue sample for tumourigenic potential. AAP
XX proteins are used to screen drugs or compounds that modulate AAP activity
XX or expression as well as treating disorders characterised by insufficient
XX or excessive production of AAP or production of AAP forms that have
XX decreased or aberrant activity compared to the wild type protein, or
```

CC modulate biological function that involve AAP. The present cDNA sequence
 CC encodes human thyroid regulated transcript (hTRG) protein which is an
 CC angiogenesis associated protein (AAP) of the invention. Human TRG is
 CC upregulated in the in vitro model of angiogenesis and is likely to be
 CC involved in signal transduction between receptors and kinases. Modulation
 CC of hTRG is useful to treat diseases related to thyroid stimulating
 CC hormone (TSH) imbalance.

SO Sequence 7506 BP; 2147 A; 1689 C; 1764 G; 1906 T; 0 other:

Query Match 99.7%; Score 496.4; DB 22; Length 7506;

Best Local Similarity 99.8%; Pred. No. 1.4e-156;

Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTTTACACCATCACCAAAACCCAGATTATATGATGATTAATATGAGTGGCCAC 60
 Db 2142 AGTTTACACCATCACCAAAACCCAGATTATATGATGATTAATATGAGTGGCCAC 2201
 QY 61 TCAGTGCATGAAAGACACCACTGTTGCTCAGATTCATGTCAGCTGTGACAACTC 120
 Db 2202 TCAGTGCATGAAAGACACCACTGTTGCTCAGATTCATGTCAGCTGTGACAACTC 2261
 QY 121 AAGTAAAGACAGACGAGAGAGGATGCTGTGAACCCAAAGTTGGCTACTCCTGGCT 180
 Db 2262 AAGTAAAGACAGACGAGAGAGGATGCTGTGAACCCAAAGTTGGCTACTCCTGGCT 2321
 QY 181 TCCCTCCCTGAAGAGGAGGTTGGTGAACAGCGACGACATCCCGTCTGGCGGAA 240
 Db 2322 TCCCTCCCTGAAGAGGAGGTTGGTGAACAGCGACGACATCCCGTCTGGCGGAA 2381
 QY 241 CCTTCTCTGGGCTATCTTGCTACCAAGAGCTGGGATGGGAGGACATATGTTCCGGA 300
 Db 2382 CCTTCTCTGGGCTATCTTGCTACCAAGAGCTGGGATGGGAGGACATATGTTCCGGA 2441
 QY 301 AATTAAATGGGTGATGAGGAGGCAAGCCACTGCTGAATAATTTCCACTCACTGTTCTAC 360
 Db 2442 AATTAAATGGGTGATGAGGAGGCAAGCCACTGCTGAATAATTTCCACTCACTGTTCTAC 2501
 QY 361 AGTGATACACGAGATCAGCATTTACATTAATTTTCCAGTACTGTCAAAAACCGAATC 420
 Db 2502 AGTGATACACGAGATCAGCATTTACATTAATTTTCCAGTACTGTCAAAAACCGAATC 2561
 QY 421 TGGAGCCCAAGCCTTAGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCCATGA 480
 Db 2562 TGGAGCCCAAGCCTTAGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCCATGA 2621
 QY 481 AGGCCACGTGATGATGCCCT 500
 Db 2622 AGGCCACGTGATGATGCCCT 2641

RESULT 10

ABK85003

ID ABK85003 standard; DNA; 6791 BP.

AC ABK85003;

DT 13-AUG-2002 (first entry)

DE DNA encoding cadherin-like asymmetry protein (CLASP) from exon 2.

KW Human: autoimmune disease; haematopoietic disorder; Digeeorge syndrome;
 KW blood protein disorder; agammaglobulinaemia; dysgammaaglobulinaemia;
 KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
 KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
 KW haemolytic anaemia; multiple sclerosis; Rheumatoid arthritis; lupus;
 KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
 KW autoimmune pulmonary inflammation; organ rejection; inflammation;
 KW CLASP; gene; ds.

OS Homo sapiens.

XX XX PN WO200231117-A2.

XX 18-APR-2002.
 PD 15-OCT-2001; 2001WO-US52202.
 XX 13-OCT-2000; 2000US-0687837.
 PR (ARBO-) ARBOR VITA CORP.
 PA (GARW/) GARMAN J D.
 PA (CAND/) CANDIA A F.
 XX Lu PS:
 DR WPI: 2002-416861/44.
 DR P-PSDB: ABG61702.
 XX
 XX New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
 PT an immune response, and for treating multiple sclerosis, rheumatoid
 PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
 PT and sepsis -
 PS Example 1: Figure 11; 245pp; English.

The invention relates to an isolated polypeptide (I) comprising an amino acid sequence that has 90 % sequence identity to one of the human cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E) sequences (PS). (I) is useful for identifying a compound or agent that binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for detecting a CLASP-2 polypeptide in a sample. (II) is useful for inhibiting a immune response in a subject. A pharmaceutical composition comprising a nucleic acid encoding (I), or (II) is useful for preventing or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where the autoimmune disease is caused or exacerbated by increased activity of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for detection or inhibition of CLASP-2 expression (e.g., antisense or ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2 polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-2 antibodies or are used as therapeutic polypeptides. The CLASP-2 polynucleotide or fragments can be used in diagnostics (e.g., as probes for CLASP-2 expression), as a lymphocyte marker and for therapeutic purposes. CLASP-2 polynucleotides can construct transgenic and knockout animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2 polynucleotides can screen for CLASP-2 agonists and antagonists. CLASP-2 polypeptides or polynucleotides can treat deficiencies or disorders of the immune system, by activating or inhibiting the activation, differentiation of immune cells and can treat or detect deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides or polynucleotides can increase differentiation and proliferation of haematopoietic cells, including the pluripotent stem cells to treat those disorders associated with a decrease in certain (or many) types of haematopoietic cells e.g., immunologic deficiency syndromes including blood protein disorders (e.g., agammaglobulinaemia, dysgammaaglobulinaemia, ataxia telangiectasia, common variable immunodeficiency, Digeeorge syndrome, lymphopenia, thrombocytopenia, or haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia, Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus, endometriosis, autoimmune thyroiditis, and autoimmune pulmonary inflammation. CLASP-2 can be used to treat anaphylaxis or hypersensitivity to an antigenic molecules, organ rejection or graft-versus-host disease (GVHD) and inflammation. ABK84922-ABK85018 CC represent cadherin-like asymmetry protein (CLASP) coding sequences and PCR primers of the invention.

SO Sequence 6791 BP; 1668 A; 1545 C; 1574 G; 1704 T; 0 other:

Query Match 99.4%; Score 496.8; DB 24; Length 6791;

Best Local Similarity 99.6%; Pred. No. 4.4e-156;

Matches 498; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTTTACACCATCACCAAAACCCAGATTATATGATGATTAATATGAGTGGCCAC 60
 Db 1985 AGTTTACACCATCACCAAAACCCAGATTATATGATGATTAATATGAGTGGCCAC 2044

QY	61	TCACCTGCATGAAAAGCAGCACCACTGTTGCTGCACATCTTCACATGTCACATGCTGTGCACACTC	120
Db	2045	TCACCTGCATGAAAAGCAGCACCACTGTTGCTGCACATCTTCACATGTCACATGCTGTGCACACTC	2104
QY	121	AAGTAAAGGAAGCAGCAGTAAGAGAGGGATGTCTGTGAACCCAAAGTTGGCTACTCCGTGCT	180
Db	2105	AAGTAAAGGAAGCAGCAGTAAGAGAGGGATGTCTGTGAACCCAAAGTTGGCTACTCCGTGCT	2164
QY	181	TCCCCTCCTGAAAGACGGAAGGGTGTGTACAGCAGCAGCACATCCGGTCTCGGCGAA	240
Db	2165	TCCCCTCCTGAAAGACGGAAGGGTGTGTACAGCAGCAGCACATCCGGTCTCGGCGAA	2224
QY	241	CCTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGAGCATTAATGTGCCGA	300
Db	2225	CCTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGAGCATTAATGTGCCGA	2284
QY	301	AATTAATAAGGTAGATGGAGGCAAGCCACTGCTGAAATAATTCACACTCATCTGTTTCTAC	360
Db	2285	AATTAATAAGGTAGATGGAGGCAAGCCACTGCTGAAATAATTCACACTCATCTGTTTCTAC	2344
QY	361	AGTGTAACTCAGGATACAGATTTACATAATTTTCCGCTACTGTGCACAAAAACCGAATC	420
Db	2345	AGTGTAACTCAGGATACAGATTTACATAATTTTCCGCTACTGTGCACAAAAACCGAATC	2404
QY	421	TGGAGCCCAAGCCTTTAGGAACACAACTGTTAAAGTACCTTAAGAGTCTGCATGACATGA	480
Db	2405	TGGAGCCCAAGCCTTTAGGAACACAACTGTTAAAGTACCTTAAGAGTCTGCATGACATGA	2464
QY	481	AGGCCACGTGATGATGGCCT	500
Db	2465	AGGCCACGTGATGATGGCCT	2484

RESULT 11
AAC87974
ID AAC87974 standard; cDNA; 4898 BP.

AAC87974;

DT 07-MAR-2001 (first entry)

Preliminary CLASP-2 nucleotide sequence #1.

KW CLASP-1: transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antinflammatory; antirheritic; cytostatic;
KW hypotensive; antihypertensive; antianemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometrios; pregnancy induced hypertension; ss.

Homo sapiens.

PN WO200061747-A2.

PD 19-OCT-2000

PF 13-APR-2000; 2000WO-US10158.

PR 14-APR-1999; 99US-0129171.

PR 14-MAY-1999; 99US-0134117.

PR 21-OCT-1999; 99US-0160860.

PR 13-DEC-1999; 99US-0170453.

14-FEB-2000; 2000US-0182296.

PA (ARBO-) ARBOR VITA CORP.

XX Lu PS;
PI

DR WPI; 2000-619230/59.

XX

PT Isolated cadherin-like asymmetry protein-2 polynucleotide and polypeptide used to diagnose, treat and prevent autoimmune diseases and

inflammatory responses - PT

PS Disclosure; Fig 10A; 286pp; English.

The present invention describes adaperlin-like asymmetry protein-2 (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory, antiinflammatory, antirheumatic, cyostatic, hypotensive, antihematic antanaemic, haemostatic and neuroprotective activities. CLASP-2 can be used to inhibit an immune response in a subject by interfering with the ability of a CLASP-2 protein to bind to another T cell or B cell. An immune response in a subject may also be inhibited by administering an antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides, proteins and antibodies can be used to prevent or treat a CLASP-2 mediated disease, such as an autoimmune disease caused or exacerbated by increased activity of TH11 cells. They can also be used to treat hypersensitivities, prevent transplantation rejection responses and augment immune responsiveness in immunodeficiency states, inhibit proliferation and differentiation of cells involved in an inflammatory response e.g. arthritis, inflammatory bowel disease and increase differentiation and proliferation of hematopoietic cells e.g. to treat anemia, thrombocytopenia and other blood protein disorders. Disorders treated by disrupting CLASP-2 function include multiple sclerosis, rheumatoid arthritis, endometriosis and pregnancy induced hypertension. The present sequence represents a prelliminary CLASP-2 nucleotide sequence, from the present invention.

Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 97.38; Score 486.4; DB 21; Length 4898;

Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY	1	AGTTTACACCATCACCCAAACCCAGAAATTTATGATGACATTTAAATPAAAGTTGCCAC	60
Db	93	AGTTTACACCATCACCCAAACCCAGAAATTTATGATGAGTTTAAATPAAAGTTGCCAC	1522
OY	61	TCAGCTGCATGAAAAAGCACCCACCTGTTGCTCACATTCCTCATGTCAGCTGTGACAATC	120
Db	153	TCAGCTGCATGAAAAAGCACCCACCTGTTGCTCACATTCCTCATGTCAGCTGTGACAATC	2112
OY	121	AAGTAAAGAGCAGCAGAAAGAGGAGTGTGTAATCCCAAGTGGGTATGTCGTGCT	180
Db	213	AAGTAAAGAGCAGCAGAAAGAGGAGTGTGTAATCCCAAGTGGGTATGTCGTGCT	272
OY	181	TCGCCCTCCTGAAAGACGGAAAGGTGTGTGACAAAGCAGACACATCCCGGCTCCGGGAA	240
Db	273	TCGCCCTCCTGAAAGACGGAAAGGTGTGTGACAAAGCAGACACATCCCGGCTCCGGGAA	3322
OY	241	CCTTCCTTCGGGGCTATCTTGGCTACCAAGACCTTGGATGGCAGGATTAATGTCGGGA	300
Db	333	CCTTCCTTCGGGGCTATCTTGGCTACCAAGACCTTGGATGGCAGGATTAATGTCGGGA	3922
OY	301	AATTAAATGGGTAGATGGAGGCAAGCCACACGTCTAAAAATTTCCACTCATCTGGTTCTAC	360
Db	393	AATTAAATGGGTAGATGGAGGCAAGCCACACGTCTAAAAATTTCCACTCATCTGGTTCTAC	452
OY	361	AGTGTATCTAGAGATCAGCATTTACATAATTTTTTCCAGTACTGTGAGAAACCGAATC	420
Db	453	AG-GGATCTACTAGAGATCAGCATTTACATAATTTTTTCCAGTACTGTGAGAAACCGAATC	511
OY	421	TGGAGCCCAACCCCTTAGAGAAACGAACCTGTGAAATTACTCTTAAGAGTGTGATGCGATGGA	480
Db	512	TGGAGCCCAACCCCTTAGAGAAACGAACCTGTGAAATTACTCTTAAGAGTGTGATGCGATGGA	571
OY	481	AGGCCACGATGATGTCGCT 500	

1;

22 pneumoniae at birth, endometritis and pregnancy induced hypertension.

matches 498; conservative 0; mismatches 1; indels 1; gaps 1;

PF 13-APR-2000; 2000WC-US10158.

```
XX 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
DR WPI: 2000-619230/59.
XX
XX Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
XX Disclosure: Fig 10C; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of Th1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
XX Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other:
SQ
Query Match 97.3%; Score 486.4; DB 21; Length 4898;
Best Local Similarity 99.6%; Pred. No. 1.2e-152;
Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 1 AGTTTACACCATCACCAAAACCCAGATTATGATGAGATTAAATGAGTTGCCAC 60
DB 93 AGTTTACACCATCACCAAAACCCAGATTATGATGAGATTAAATGAGTTGCCAC 152
OY 61 TCAGCTGCATGAAAGACACCACTGTTGCTCATCTTCATGTCAGTGTGACAACAC 120
DB 153 TCAGCTGCATGAAAGACACCACTGTTGCTCATCTTCATGTCAGTGTGACAACAC 212
OY 121 AAGTTAAGACGACGACGACGATGCTGTTGAACCCAAAGTTGGTACTCTGGCT 180
DB 213 AAGTTAAGACGACGACGACGATGCTGTTGAACCCAAAGTTGGTACTCTGGCT 272
OY 181 TCCCTCTCTGAAAGACGAGAGGTGTGACAAGCAGCAGCATCCGGTCTGGCGAA 240
DB 273 TCCCTCTCTGAAAGACGAGAGGTGTGACAAGCAGCAGCATCCGGTCTGGCGAA 332
OY 241 CCTTCCCTGGGCTATCTGGTACCAAGAGCTGGAGTGGGAGGACATTAATGTCGGA 300
DB 333 CCTTCCCTGGGCTATCTGGTACCAAGAGCTGGAGTGGGAGGACATTAATGTCGGA 392
A 301 AATTAATGGGTGATGAGGACGACGACCTGCTGAATAATTTCCACTCATCTGTTTCTAC 360
OY 301 AATTAATGGGTGATGAGGACGACGACCTGCTGAATAATTTCCACTCATCTGTTTCTAC 360
```

```
DB 393 AATTAATGGGTGATGAGGACGACGACCTGCTGAATAATTTCCACTCATCTGTTTCTAC 452
OY 361 AGTGTATCTACAGCATCAGCATTTACATAATTTTTCACACTACTGTCAGAAACCAATC 420
DB 453 AG-GGATACCTACAGCATCAGCATTTACATAATTTTTCACACTACTGTCAGAAACCAATC 511
OY 421 TGGAGCCCAAGCCTTACGAAACGAACTTTGAAGTACCTTAAGAGTCTGCATCGATGCA 480
DB 512 TGGAGCCCAAGCCTTACGAAACGAACTTTGAAGTACCTTAAGAGTCTGCATCGATGCA 571
OY 481 AGGCCACGTGATGATCGCCT 500
DB 572 AGGCCACGTGATGATCGCCT 591
RESULT 14
AAC87977
ID AAC87977 standard; cDNA; 4898 BP.
XX
AC AAC87977:
XX
XX 07-MAR-2001 (first entry)
XX
XX Preliminary CLASP-2 nucleotide sequence #4.
XX
XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
XX Homo sapiens.
XX
XX WO200061747-A2.
XX
XX 19-OCT-2000.
XX
XX 13-APR-2000; 2000WO-US10158.
XX
XX 14-APR-1999; 99US-0129171.
XX 14-MAY-1999; 99US-0134114.
XX 14-MAY-1999; 99US-0134117.
XX 14-MAY-1999; 99US-0134118.
XX 21-OCT-1999; 99US-0160860.
XX 29-OCT-1999; 99US-0162498.
XX 13-DEC-1999; 99US-0170453.
XX 14-JAN-2000; 2000US-0176195.
XX 14-FEB-2000; 2000US-0182296.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX Lu PS;
XX
XX WPI: 2000-619230/59.
XX
XX Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
XX Disclosure: Fig 10D; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
```



```
OY 181 TCCCTCTCTGAAAGAGGAGGGTGTGACAAAGCAGCAGACATCCCGTCTCGCGGAA 240
    |||||||
Db 273 TCCCTCTCTGAAAGAGGAGGGTGTGACAAAGCAGCAGCAGACATCCCGTCTCGCGGAA 332
    |||||||
OY 241 CCTTCTTCGGGCTATCTTGGCTACCAAGAGCTTGGATGGGCAAGGCAATTATGGTCCGGA 300
    |||||||
Db 333 CCTTCTTCGGGCTATCTTGGCTACCAAGAGCTTGGATGGGCAAGGCAATTATGGTCCGGA 392
    |||||||
OY 301 AATTAAATGGGTAGATGAGGCAAGCAGCTGCTGAAATTTCCACTCATCTGTTTCTTAC 360
    |||||||
Db 393 AATTAAATGGGTAGATGAGGCAAGCAGCTGCTGAAATTTCCACTCATCTGTTTCTTAC 452
    |||||||
OY 361 AGTGTATCTCAGATCAGCATTTTACATTAATTTTCCAGTACTGTCCAGAAAACCGAATC 420
    || || |||||||
Db 453 AG-GGATACTCAGATCAGCATTTTACATTAATTTTCCAGTACTGTCCAGAAAACCGAATC 511
    || || |||||||
OY 421 TGGAGGCCAAGGCTTAGGAAGCACTGTAAAGTACCTTAAGAGTCTGCATGGCATGGA 480
    |||||||
Db 512 TGGAGGCCAAGGCTTAGGAAGCACTGTAAAGTACCTTAAGAGTCTGCATGGCATGGA 571
    |||||||
OY 481 AGGCCACGTGATGATCGCCT 500
    |||||||
Db 572 AGGCCACGTGATGATCGCCT 591
    |||||||
```

Search completed: February 7, 2003, 07:07:37
-Job time : 137.83 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:00:34 ; Search time 28.6285 Seconds
(without alignments)
5356.145 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500

Perfect score: 500
Sequence: 1 agtttacaccatcaccaaa.....agccacgtgatgcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30.4	6.1	1005	US-08-936-165A-228	Sequence 228, App
C 2	30.4	6.1	1920	US-09-160-496-4	Sequence 4, Appl
C 3	30.2	6.0	1972	US-08-961-083-203	Sequence 203, App
C 4	30.2	6.0	14231	US-08-961-527-81	Sequence 81, Appl
C 5	30	6.0	2520	US-08-405-254-9	Sequence 9, Appl
C 6	29.4	5.9	442	US-09-071-035-443	Sequence 443, App
C 7	29.4	5.9	606	US-09-071-035-441	Sequence 441, App
C 8	29.4	5.9	1006	US-08-924-747-15	Sequence 15, Appl
C 9	29.4	5.9	1006	US-09-247-373B-15	Sequence 15, Appl
C 10	29.4	5.8	726	US-08-660-645A-11	Sequence 11, Appl
C 11	29.2	5.8	726	US-09-298-718-11	Sequence 11, Appl
C 12	29.2	5.8	726	US-09-546-968-11	Sequence 11, Appl
C 13	29.2	5.8	726	US-08-980-832-28	Sequence 28, Appl
C 14	29.2	5.8	726	US-08-980-832-28	Sequence 28, Appl
C 15	29.2	5.8	4403765	US-09-103-840A-2	Sequence 2, Appl
C 16	29.2	5.8	4411529	US-09-103-840A-1	Sequence 1, Appl
C 17	29	5.8	856	US-08-117-373-10	Sequence 10, Appl
C 18	29	5.8	1542	US-09-008-271A-13	Sequence 13, Appl
C 19	29	5.8	4084	US-08-568-459A-1	Sequence 1, Appl
C 20	29	5.8	4084	US-08-487-826B-1	Sequence 1, Appl
C 21	29	5.8	4084	US-09-210-288-1	Sequence 1, Appl
C 22	29	5.8	4084	US-09-210-288-1	Sequence 1, Appl
C 23	29	5.8	4156	US-08-961-527-211	Sequence 211, App
C 24	28.8	5.8	1747	US-07-808-455A-2	Sequence 2, Appl
C 25	28.6	5.7	1643	US-08-383-750-3	Sequence 3, Appl
C 26	28.6	5.7	1643	US-08-352-678-3	Sequence 3, Appl
C 27	28.6	5.7	1643	PCT-US93-09636-3	Sequence 3, Appl

28	28.4	5.7	1072	1	US-07-971-096-1	Sequence 1, Appl
29	28.4	5.7	1072	1	US-08-175-096-1	Sequence 1, Appl
30	28.2	5.6	608	4	US-09-385-982-236	Sequence 236, App
C 31	28.2	5.6	18609	4	US-08-943-731-1	Sequence 1, Appl
C 32	28	5.6	488	4	US-09-385-982-368	Sequence 368, App
C 33	28	5.6	111282	4	US-09-754-250-3	Sequence 3, Appl
C 34	27.8	5.6	2970	4	US-09-110-517-3	Sequence 90, Appl
C 35	27.8	5.6	3826	4	US-09-302-620B-90	Sequence 1, Appl
C 36	27.8	5.6	9412	2	US-08-955-138-1	Sequence 1, Appl
C 37	27.6	5.5	1463	4	US-09-399-913-1	Sequence 1, Appl
C 38	27.6	5.5	1463	4	US-09-288-731-1	Sequence 1, Appl
C 39	27.6	5.5	1540	4	US-09-399-913-9	Sequence 9, Appl
C 40	27.6	5.5	1540	4	US-09-298-731-9	Sequence 9, Appl
C 41	27.6	5.5	1907	4	US-09-399-913-5	Sequence 5, Appl
C 42	27.6	5.5	1907	4	US-09-298-731-5	Sequence 5, Appl
C 43	27.6	5.5	2624	1	US-08-032-382B-1	Sequence 1, Appl
C 44	27.4	5.5	518	4	US-09-123-912-87	Sequence 87, Appl
C 45	27.4	5.5	518	4	US-09-643-597-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
: Sequence 228, Application US/08936165A
: Patent No. 6348582
: GENERAL INFORMATION:
: APPLICANT: Black, Michael
: APPLICANT: Burnham, Martin
: APPLICANT: Hodgson, John
: APPLICANT: Knowles, David
: APPLICANT: Lonetto, Michael
: APPLICANT: Nicholas, Richard
: APPLICANT: Pratt, Julie
: APPLICANT: Reichard, Richard
: APPLICANT: Rosenberg, Martin
: APPLICANT: Ward, Judith
: TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
: TITLE OF INVENTION: Polypeptides and Their Uses
: NUMBER OF SEQUENCES: 534
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/936,165A
: FILING DATE: 24-SEP-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/027,032
: FILING DATE: 24-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Gimmil, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P50549
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
: TELE: :
: INFORMATION FOR SEQ ID NO: 228:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1005 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single

```

;
;   TOPOLOGY: linear
;   MOLECULE TYPE: Genomic DNA
US-08-936-165A-228

Query Match
Best Local Similarity 6.1%; Score 30.4; DB 4; Length 1005;
Matches 65; Conservative 0; Mismatches 36; Indels 3; Gaps 1;

QY 2 GTTTTACCATCACCACCAAAACCCAGAAATTTATGA---TGAGATTAAATAGAGTTGCC 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 439 GTTGTCACCAATATATATATATATATATATATATATATATATATATATATATATATAT 380

QY 59 ACTACAGCTGCATGAAAGACACACCACTGTGTCATCATTTCTTCCA 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 379 ACTCCGCTGCATCTAGACACCAATCTTCCACTGTACTCTGA 336

RESULT 2
US-09-160-496-4/C
; Sequence 4, Application US/09160496
; Patent No. 6346613
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J
; TITLE OF INVENTION: Composition and Method for Enhancing Paracellular
; FILE REFERENCE: Docket No. 6346613; 98,1070.US
; CURRENT APPLICATION NUMBER: US/09/160,496
; EARLIER FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: US 60/059,644
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 4
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Gallus gallus
; PUBLICATON INFORMATION:
; TITLE: Occludin: A novel integral membrane protein localizing
; TITLE: at tight junctions
; JOURNAL: J. Cell Biol.
; VOLUME: 123
; ISSUE: 6
; PAGES: 1777-1788
; DATE: Dec 1993
; DATABASE ACCESSION NUMBER: D21837
US-09-160-496-4

Query Match
Best Local Similarity 6.1%; Score 30.4; DB 4; Length 1920;
Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 182 CCCCTCCTGAAGAGGAAAGGTGTGACAGGACGACACATCCCGGCTCGGGGAGAC 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 842 CCCCTCCTGACACGCGGCGGCGGTCCAGTAGATGTTGCTTCCGTAGGCGCAGAT 783

QY 242 CTTTCCTGGGGCTATCTTGCTACCAAGAGCTTGGAGTGGCAGCATATATGTCGGAA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 782 CTTACTCGCGCTCTTCTGGGCGAAGAGAGATGAGGACAGCAGATGATGATGAGAA 723

RESULT 3
US-08-961-083-203
; Sequence 203, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
```

```

;
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII text
;
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,083
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 203:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1972 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
US-08-961-083-203

Query Match
Best Local Similarity 6.0%; Score 30.2; DB 3; Length 1972;
Matches 62; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 238 GAACCTCTTCGGGCTATCTTGCTACCAAGAGCTTGGATGGGACGATTATGCTCC 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 44 GAATGCCATTGTGCTTCTCCGTGTATGCAATGATATATGTAACGGATGTATGTTGG 103

QY 298 GGAATTTAATGGTAGATGAGGACCACTGCTGAAATTTCCACTACTCTG 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 104 GGGAAACACATTGAAGGAGCATGGAAACCTATCTCTGAAGATGTAAGCCATTAG 158

RESULT 4
US-08-961-527-81
; Sequence 81, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII text
;
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,527
;   FILING DATE:
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
```

```

; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-81

Query Match
Best Local Similarity 53.9%; Score 30.2; DB 4; Length 14231;
Pred. No. 11;
Matches 62; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 238 GAACCTTCCTGGGCGATCTTGCTACCAAGCTTGGATGGGAGGATATATGTC 297
Db 9070 GAATCGGCATTGTTGCTTCTCCGTGTATGCAATGATTAATGTTAAGGATATGTCG 9129

QY 298 GGAATTAAATGGGTAGATGAGGCAAGCCACTGCTGAAATTTCCATCATCTG 352
Db 9130 GGGAAACACATTGGAAGGAGCATGGGAAGCTATTCCTGAGAGTGTAAAGCCATATG 9184

RESULT 5
US-08-405-254-9/C
; Sequence 9, Application US/08405254
; Patent No. 5773288
; GENERAL INFORMATION:
; APPLICANT: BRIGGS, STEVEN P.
; APPLICANT: BENSON, ROBERT J.
; TITLE OF INVENTION: PLANT GENES AFFECTING GIBBERELLIC ACID
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K STREET, NW, SUITE 500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,254
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/261,465
; FILING DATE: 17-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, STEPHEN A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/299/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-405-254-9

Query Match
Best Local Similarity 47.4%; Score 30; DB 1; Length 2520;
Pred. No. 5;
Matches 90; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```

```

QY 149 GTCTGTAACCCAGTGGCTACTCTGCGCTTCCCTCGTGAAGAGGAGGTGTC 208
Db 2219 GTTGTCTGTCGCCGTGGCCGGGSCCTGCTGCTGCTCCACACCCGGCAGCTTAGG 2160

QY 209 ACAAGGAGCAGACACATCCCGGCTCTGGCGAACCCTTCTTGGGGTATCTTGGTACCA 268
Db 2159 CTGCCGGGGGGCCCGTGCGCCGCCGCTCCTCCTCCTGCTGCTGCTTTGCCA 2100

QY 269 GAGCTTGGGATGGCAGGCAATATGTCGCCGAATTAATGGTATGATGAGCAAGCCA 328
Db 2099 CGGATACGATGCGGATAGAGCGCATGACCTGATGACGAGCTGACGCAAGTGA 2040

QY 329 CTGCTGAAA 338
Db 2039 AAGCAGAAAA 2030

RESULT 6
US-09-071-035-443
; Sequence 443, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 443:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 442 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-071-035-443

Query Match
Best Local Similarity 5.9%; Score 29.4; DB 4; Length 442;
Pred. No. 3;
Matches 81; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 300 AATTAATGGGTAGATGAGGCAAGCCACTGTAATAATTTCCACTCATCTGGTTCTA 359
Db 240 AAAAAAATGAATGATGATGTTAAACTGAGTAGCTTAGTTCCGTAGGAAAGTTACAA 299

QY 360 CAGTGTATACAGAGATCAGCATTTACATTAATTTTTCAGTACTGTACAGAAACCGAAT 419
Db 300 CGATTTAATAAATGATGACATGATGATATGCTTACTTTGGATTGGAACAAATGGA 359

QY 420 CTGGAGCCCAAGCCTTAGAACAAGCACTTGTAAAGTACCTTAAGAGT 466
```

DB 360 CGCAGACGAGTGTACTTTGGCTAAGGTGAAGAACATTAATT 406

RESULT 7
US-09-071-035-441

; Sequence 441, Application US/09071035

; Patent No. 6448043

; GENERAL INFORMATION:

; APPLICANT: Gil H. Choi

; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

; NUMBER OF SEQUENCES: 496

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/071,035

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: A. Anders Brooks

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB369P2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 441:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 606 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; US-09-071-035-441

Query Match 5.9%; Score 29.4; DB 4; Length 606;
Best Local Similarity 48.5%; Pred. No. 3.6;
Matches 81; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 300 AATTAATGGTAGATGAGGAGCAGCCACTGCTGAAATTTCCACTCATCTGTTCTA 359

DB 401 AAAAATGGAATGATGATGAACTGAGTACGTTAGTTCCGTAGAGAAAGTTACAA 460

QY 360 CAGGTGATCTGAGATGAGTATTTCCAGTACCTGAGTACGAAACGAAAT 419

DB 461 CGATTTTAAAAATGATCAACATGATATGCTTTACTTTTGGATTGACAAATGCAA 520

QY 420 CTGAGGCCCAAGCCTTAGAAGAACTGTAAGTACTTAAGT 466

DB 521 CGCAGACGAGGTGTACTTTGGCTAAGGTGAAGAACATTAATT 567

RESULT 8

US-08-924-747-15/C

; Sequence 15, Application US/08924747

; Patent No. 6063570

; GENERAL INFORMATION:

; APPLICANT: MCGONIGLE, BRIAN

; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95

; SOFTWARE: MICROSOFT WORD VERSION 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/924,747

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: FLOYD, LINDA AXAMETHY

; REGISTRATION NUMBER: 33,692

; REFERENCE/DOCKET NUMBER: CL-1108

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-773-0164

; TELEFAX: 302-892-8112

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1006 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; TISSUE TYPE: SOYBEAN

; IMMEDIATE SOURCE:

; CLONE: SRL.PK0011.D6

; US-08-924-747-15

Query Match 5.9%; Score 29.4; DB 3; Length 1006;
Best Local Similarity 60.8%; Pred. No. 4.8;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 63 AGCTGATGAAGAGCAGCCGCTGCTCAGATCTTCATGCTGAGTGAACATCAA 122

DB 412 ACCTCAATTCCTCTCAACATCTTCTCAGCGCTTCTCATCAGCTGTAAGAAACGAT 353

QY 123 GTAAGAGAGCAGCAAGAA 141

DB 352 TTCATGCGAGCAGCAAAA 334

RESULT 9

US-09-247-373B-15/C

; Sequence 15, Application US/09247373B

; Patent No. 6168954

; GENERAL INFORMATION:

; APPLICANT: MCGONIGLE, BRIAN

; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES

; FILE REFERENCE: CL-1108-A

; CURRENT APPLICATION NUMBER: US/09/247,373B

; PRIOR FILING DATE: 1999-02-10

; PRIOR APPLICATION NUMBER: 08/924,747

; PRIOR FILING DATE: 1997-09-05

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 15

; LENGTH: 1006

; TYPE: DNA

; ORGANISM: SOYBEAN

US-09-247-373B-15

Query Match 5.9%; Score 29.4; DB 4; Length 1006;

Best Local Similarity 60.8%; Pred. No. 4.8;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 63 AGCTGATGAAAGCAGCAGCTGTGCTCATCTTTCATGTCAGCTGTGACCACTCA 122

Db 412 AGCTCAATTCCTCTCCACATCTTCTCAGCTCTTTCATCAGCTGTGAAACGGAT 353

QY 123 GTAAGGAGCAGCAGAA 141

Db 352 TTCATGACGACCAAAAA 334

RESULT 10

US-09-296-715-15/c

; Sequence 15, Application US/09296715

; Patent No. 6171839

; GENERAL INFORMATION:

; APPLICANT: MCGONIGLE, BRIAN

; APPLICANT: O'KEEFE, DANIEL

; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95

; SOFTWARE: MICROSOFT WORD VERSION 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/296,715

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: FLOYD, LINDA AXAMETHY

; REGISTRATION NUMBER: 33,692

; REFERENCE/DOCKET NUMBER: CL-1108

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-892-8112

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1006 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHEetical: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; TISSUE TYPE: SOYBEAN

; IMMEDIATE SOURCE:

; CLONE: SRL.PK0011.D6

US-09-296-715-15

Query Match

Best Local Similarity 5.9%; Score 29.4; DB 4; Length 1006;

Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 63 AGCTGATGAAAGCAGCAGCTGTGCTCATCTTTCATGTCAGCTGTGACCACTCA 122

Db 412 AGCTCAATTCCTCTCCACATCTTCTCAGCTCTTTCATCAGCTGTGAAACGGAT 353

QY 123 GTAAGGAGCAGCAGAA 141

Db 352 TTCATGACGACCAAAAA 334

RESULT 11

US-08-660-645A-11/c

; Sequence 11, Application US/08660645A

; Patent No. 6087152

; GENERAL INFORMATION:

; APPLICANT: Hohmann, Hans-Peter

; APPLICANT: Pasamontes, Luis

; APPLICANT: Tessier, Michel

; APPLICANT: van Loon, Adolphus

; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION

; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: NJ

; COUNTRY: USA

; ZIP: 07110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentln Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/660,645A

; FILING DATE: 07-JUN-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 95108888.9

; FILING DATE: 09-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Pokras, Bruce A.

; REGISTRATION NUMBER: 32,748

; REFERENCE/DOCKET NUMBER: RAN 6002/170

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-5801

; TELEFAX: (201) 235-2363

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 726 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-660-645A-11

Query Match

Best Local Similarity 5.8%; Score 29.2; DB 3; Length 726;

Matches 82; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 276 GGATGGCAGCATTTATGTCGGAAATTAATGGTATGATGAGGCAAGCCACTGCTGA 335

Db 557 GGACGGTGGCGCAGCAGCAGTACGAAAGCAAGATCTGGAGGAAAGCAGCAAGCGGA 498

QY 336 AAATTCACATCTGTTCTTCTACAGTATCTACAGATCAGATTCATATTTT 395

Db 497 AOCGGCCAGAGTATACATTCACAGCGTACCCAGAGATCAAGGTAGTACG 438

QY 396 TCCACTACTGTAGAAACGAAATCTGGAGCCCAAGCCTTAGAAACGAA 445

Db 437 ATAACGGCAGCAGCAGACCTTACAGCCCAAGCAAGTGAAGAAACGAA 388

RESULT 12

US-09-298-718-11/c

; Sequence 11, Application US/09298718

; Patent No. 6124113

; GENERAL INFORMATION:

; APPLICANT: Hohmann, Hans-Peter

; APPLICANT: Pasamontes, Luis

; APPLICANT: Tessier, Michel

; APPLICANT: van Loon, Adolphus

; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION

```

: NUMBER OF SEQUENCES: 47
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street
: CITY: Nutley
: STATE: NJ
: COUNTRY: USA
: ZIP: 07110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/298,718
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/660,645
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Pokras, Bruce A.
: REGISTRATION NUMBER: 32,748
: REFERENCE/DOCKET NUMBER: RAN 6002/170
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (201) 235-5801
: TELEFAX: (201) 235-2363
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 726 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-09-298-718-11

Query Match          5.8%; Score 29.2; DB 3; Length 726;
Best Local Similarity 48.2%; Pred. No. 4.6;
Matches 82; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 276 GGATGGGCGAGCATTTATGCTCCGGAATTAAATGGGTAGATGGAGCGACCACTGCTGA 335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 557 GGACGGTGGCGGACGACGATACCGAAACGAAGATCTGGATGGAACCCAGAACGCCGA 498

QY 336 AATTTCACACTCATCTGCTTTCACAGTGTATCTAGATCAGCATTTACATATTTT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 497 ACCGGCCAGAGATTAACGTACATCCACGGTCCACCGAGATCAGACCGGTAGGTACG 438

QY 396 TCCAGTACTGTCTCAAAACCGAATCTGAGCCCAACCTTAGGAACGAA 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 ATTAACGGCGACGACGACCTTCACGCCCAACGAGTAGTGGAAACGAA 388

RESULT 13
US-09-546-969-11/c
: Sequence 11, Application US/09546969
: Patent No. 6207409
: GENERAL INFORMATION:
: APPLICANT: Hoffmann, Hans-Peter
: APPLICANT: Pasamontes, Luis
: APPLICANT: Tessier, Michel
: APPLICANT: van Loon, Adolphus
: TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
: NUMBER OF SEQUENCES: 47
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street
: CITY: Nutley
: STATE: NJ
: COUNTRY: USA
: ZIP: 07110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/546,969
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/660,645
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Pokras, Bruce A.
: REGISTRATION NUMBER: 32,748
: REFERENCE/DOCKET NUMBER: RAN 6002/170
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (201) 235-5801
: TELEFAX: (201) 235-2363
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 726 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-09-546-969-11

Query Match          5.8%; Score 29.2; DB 4; Length 726;
Best Local Similarity 48.2%; Pred. No. 4.6;
Matches 82; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 276 GGATGGGCGAGCATTTATGCTCCGGAATTAAATGGGTAGATGGAGCGACCACTGCTGA 335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 557 GGACGGTGGCGGACGACGATACCGAAACGAAGATCTGGATGGAACCCAGAACGCCGA 498

QY 336 AATTTCACACTCATCTGCTTTCACAGTGTATCTAGATCAGCATTTACATATTTT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 497 ACCGGCCAGAGATTAACGTACATCCACGGTCCACCGAGATCAGACCGGTAGGTACG 438

QY 396 TCCAGTACTGTCTCAAAACCGAATCTGAGCCCAACCTTAGGAACGAA 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 ATTAACGGCGACGACGACCTTCACGCCCAACGAGTAGTGGAAACGAA 388

RESULT 14
US-08-980-832-28/c
: Sequence 28, Application US/08980832B
: Patent No. 6291204
: GENERAL INFORMATION:
: APPLICANT: Pasamontes, Luis
: APPLICANT: Tsygankov, Yuri
: TITLE OF INVENTION: Improved Fermentative Carotenoid Production
: FILE REFERENCE: Improved Fermentative Carotenoid
: CURRENT APPLICATION NUMBER: US/08/980,832B
: CURRENT FILING DATE: 1997-12-01
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 28
: LENGTH: 726
: TYPE: DNA
: ORGANISM: Alcaligenes PC-1
: US-08-980-832-28

Query Match          5.8%; Score 29.2; DB 4; Length 726;
Best Local Similarity 48.2%; Pred. No. 4.6;
Matches 82; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 276 GGATGGGCGAGCATTTATGCTCCGGAATTAAATGGGTAGATGGAGCGACCACTGCTGA 335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 557 GGACGGTGGCGGACGACGATACCGAAACGAAGATCTGGATGGAACCCAGAACGCCGA 498

QY 336 AATTTCACACTCATCTGCTTTCACAGTGTATCTAGATCAGCATTTACATATTTT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 497 ACCGGCCAGAGATTAACGTACATCCACGGTCCACCGAGATCAGACCGGTAGGTACG 438
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:01:14 ; Search time 32.2903 Seconds
(without alignments)
7339.716 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agttaccacatcaccaaa.....aggccacgtgatgcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204.4	40.9	6454	10	US-09-736-969A-1
2	85	17.0	4391	10	US-09-736-969A-7
3	71.4	14.3	7215	10	US-09-736-960-1
4	46	9.2	6372	9	US-09-736-968A-1
5	40.4	8.1	211	9	US-09-736-968A-77
6	34.6	6.9	6691	10	US-09-070-927A-88
7	32.4	6.5	1314	9	US-09-738-626-795
8	32.4	6.5	3309400	9	US-09-738-626-1
9	31.8	6.4	659158	9	US-09-771-208-20
10	31.6	6.3	14141	10	US-09-070-927A-394
11	31.4	6.3	4915	10	US-09-070-927A-125
12	31.2	6.2	2000	9	US-09-938-842A-1143
13	30.8	6.2	417	10	US-09-960-352-12649
14	30.6	6.1	32191	10	US-09-764-864-1678
15	30.4	6.1	1005	10	US-09-939-980-228
16	30.4	6.1	326014	10	US-09-731-231A-3
17	30.2	6.0	477	10	US-09-764-877-623
18	30.2	6.0	1972	10	US-09-765-272-203
19	29.6	5.9	405	10	US-09-960-352-2213

C 20	29.4	5.9	251	9	US-09-796-692-9770	Sequence 9270, Ap
C 21	29.4	5.9	371	10	US-09-864-761-32600	Sequence 32600, A
C 22	29.4	5.9	528	9	US-09-796-652-3999	Sequence 3999, Ap
C 23	29.4	5.9	1686	9	US-10-114-893-37	Sequence 37, Appl
C 24	29.4	5.9	2707	10	US-09-917-800A-1420	Sequence 1420, Ap
C 25	29.4	5.9	7925	10	US-09-070-927A-311	Sequence 311, App
C 26	29.4	5.9	9450	9	US-10-126-912-2	Sequence 2, Appl1
C 27	29.4	5.9	9457	9	US-10-126-912-7	Sequence 7, Appl1
C 28	29.4	5.9	9828	9	US-10-126-912-1	Sequence 1, Appl1
C 29	29.2	5.8	726	9	US-09-920-923-28	Sequence 28, Appl
C 30	29.2	5.8	726	10	US-09-547-267-11	Sequence 11, Appl
C 31	29.2	5.8	1685	10	US-09-827-998-19	Sequence 11, Appl
C 32	29.2	5.8	1944	10	US-09-864-761-2825	Sequence 2825, Ap
C 33	29.2	5.8	4158	10	US-09-864-761-2825	Sequence 2825, Ap
C 34	29.2	5.8	5313	10	US-09-827-998-15	Sequence 15, Appl
C 35	29.2	5.8	5376	10	US-09-827-998-2	Sequence 2, Appl1
C 36	29.2	5.8	5461	10	US-09-827-998-8	Sequence 8, Appl1
C 37	29.2	5.8	6719	10	US-09-827-998-1	Sequence 1, Appl1
C 38	29	5.8	416	9	US-09-736-457-469	Sequence 469, App
C 39	29	5.8	416	9	US-09-902-941-469	Sequence 469, App
C 40	29	5.8	416	9	US-09-848-626-469	Sequence 469, App
C 41	29	5.8	494	10	US-09-920-300A-446	Sequence 446, App
C 42	29	5.8	494	12	US-10-033-528-446	Sequence 446, App
C 43	29	5.8	590	10	US-09-864-761-9899	Sequence 9899, Ap
C 44	29	5.8	954	10	US-09-947-971-3	Sequence 3, Appl1
C 45	29	5.8	2586	10	US-09-804-551B-19	Sequence 19, Appl1

ALIGNMENTS

RESULT 1
US-09-736-969A-1
Sequence 1, Application US/09736969A
Patent No. US20020068302A1
GENERAL INFORMATION:
APPLICANT: Lu, Peter
APPLICANT: Garman, Jonathan David
APPLICANT: Candia III, Albert Frederick
APPLICANT: Arbor Vita Corporation
TITLE OF INVENTION: CLASP-4 Transmembrane Protein
FILE REFERENCE: 020054-000411US
CURRENT APPLICATION NUMBER: US/09/736, 969A
CURRENT FILING DATE: 2000-12-13
PRIOR APPLICATION NUMBER: US 60/160, 860
PRIOR FILING DATE: 1999-10-21
PRIOR APPLICATION NUMBER: US 60/162, 498
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: US 60/170, 453
PRIOR FILING DATE: 1999-12-13
PRIOR APPLICATION NUMBER: US 60/176, 195
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: US 60/182, 296
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: US 09/547, 276
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196, 267
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196, 460
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196, 527
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196, 528
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 09/687, 837
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240, 503
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240, 508
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240, 539
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240, 543

;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 153
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 1
;; LENGTH: 6454
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: full length human CLASP-4 cDNA
;; NAME/KEY: CDS
;; LOCATION: (95)..(6121)
;; OTHER INFORMATION: human CLASP-4
US-09-736-969A-1

Query Match 40.9%; Score 204.4; DB 10; Length 6454;
Best Local Similarity 63.4%; Pred. No. 8.1e-57;
Matches 313; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 2 GTTTACACATCACCAACCAGATTTTATGATGATTAATAAGTGGCCACT 61
DB 2000 GTCTCCATCACACCAATATCCAGAGTTCTATGATGAGATTAAATGAGCTTCCACT 2059
QY 62 CAGCTGCATGAAAAGCACCCTGTGTGCATCTTCCATGTCAGCTGACACTCA 121
DB 2060 CACCTCATCAAAAACATCATTTGCTTTCACCTTTTATCATGTAAGTTGTAATTAAC 2119
QY 122 AGTAAGAAGCAGCAAGAGGAGTGTGTAAGCCCAAGTTGCTACTCTGGCTT 181
DB 2120 ACNAGGAAACCAACCAAGAGACAGATTAAGTGAACCTCAGTTGGGTTGCCGCGTA 2179
QY 182 CCCCCTCTGAAGAGGAGGGGTGACAGGAGCAGACATCCGGCTCCGGCGAAC 241
DB 2180 CCTTCTGTAAGATGTTGATATCATCATTTTGAACAGACTGCGCACTTCCGCCAAT 2239
QY 242 CTTCCTTGGGCTATCTTGGCTACCAAGAGCTTGGAGTGGCAGCATTAATGATCGGAA 301
DB 2240 CTTCCTCCAGGCTACTTGAATGATGATGAGATCAAGAGGAATGTAAGTGAT 2299
QY 302 ATTAATGGGTAGATGAGGCAAGCCACTGCTGAATAATTTCCATCATCTGTTCTACA 361
DB 2300 ATTAATGGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2359
QY 362 GTGTATAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
DB 2360 ATTTAAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2419
QY 422 GGAGCCCAAGCCTTAGAAGCAAGCACTTGTAAAGTACCTTAAGAGTCTGCATGCGA 481
DB 2420 GGCTCGAAGAGAGTTCAGAGGAGCTCATTAATAATTAAAGTGTTCATGCCATGGAG 2479
QY 482 GGCACAGTATGAT 495
DB 2480 ATCCAAGTATGAT 2493

RESULT 2
US-09-736-969A-7
; Sequence 7, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736,969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29

;; PRIOR APPLICATION NUMBER: US 60/170,453
;; PRIOR FILING DATE: 1999-12-13
;; PRIOR APPLICATION NUMBER: US 60/176,195
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 60/182,296
;; PRIOR FILING DATE: 2000-02-14
;; PRIOR APPLICATION NUMBER: US 09/547,276
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,267
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,460
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,527
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,528
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 09/687,837
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,503
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,508
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,539
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,543
;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 153
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 7
;; LENGTH: 4391
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: human CLASP-4 cDNA
;; NAME/KEY: CDS
;; LOCATION: (414)..(4058)
;; OTHER INFORMATION: human CLASP-4
US-09-736-969A-7

Query Match 17.0%; Score 85; DB 10; Length 4391;
Best Local Similarity 52.4%; Pred. No. 1.5e-17;
Matches 187; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 140 AAGAGGATGCTGTGAACCCAGTGGCTACCTCCGGCTCCGCCCTGAAGAAGCA 199
DB 78 AAGCGGGCCCCGGTGGACCTCCGGTGGGTTGCTGGGTCCTTGTGGAAGTTGGT 137
QY 200 AAGGTGTGACNACGAGACACATCCGGTCTCGGCGAACCCTTCTCGGGCTATCTT 259
DB 138 AGATTCCTTCCCTTTTGGGGGCGGCTCGCGSTTCCGCCAATCTTCCCGGGGCTCCTTA 197
QY 260 GGTACCAAGAGCTTGGAGTGGCAGCATTAATGCTCCGGAATTAATGATGATGCA 319
DB 198 ATTCGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 257
QY 320 GGCAAGCCACTGCGAATAATTCACATCTGCTGCTCTACAGTATACAGATGATGAT 379
DB 258 GTAAGCCCTTGTGAGAGTTAAAGGCCCTTGGGTCTTCCATTTTCCCTTAAGTTCTG 317
QY 380 CATTTACATAATTTTTCACATGCTAGCAGNAAACCAATCTGGAGCCCAAGCCTTAGGA 439
DB 318 CATTTGCCCAATATCTTCCATCATTTGGCGGCTGATTTGGTGACAGCTGGAAGAGTTCCA 377
QY 440 AAGCACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAAGGCCACGATGATC 496
DB 378 GGGGGGCTCATTAATTTTAAAGTCTTTCATGATGATGATGATGATGATGATGATGATGAT 434

RESULT 3
US-09-736-960-1
; Sequence 1, Application US/09736960
; Patent No. US20020102267A1

```

? GENERAL INFORMATION:
? APPLICANT: Lu, Peter
? APPLICANT: Garman, Jonathan David
? APPLICANT: Candia III, Albert Frederick
? APPLICANT: Arbor Vita Corporation
? TITLE OF INVENTION: CLASP-5 Transmembrane Protein
? FILE REFERENCE: 020054-000511US
? CURRENT APPLICATION NUMBER: US/09/736,960
? CURRENT FILING DATE: 2001-09-20
? PRIOR APPLICATION NUMBER: US 60/160,860
? PRIOR FILING DATE: 1999-10-21
? PRIOR APPLICATION NUMBER: US 60/162,498
? PRIOR FILING DATE: 1999-10-29
? PRIOR APPLICATION NUMBER: US 60/170,453
? PRIOR FILING DATE: 1999-12-13
? PRIOR APPLICATION NUMBER: US 60/176,195
? PRIOR FILING DATE: 2000-01-14
? PRIOR APPLICATION NUMBER: US 60/182,296
? PRIOR FILING DATE: 2000-02-14
? PRIOR APPLICATION NUMBER: US 09/547,276
? PRIOR FILING DATE: 2000-04-11
? PRIOR APPLICATION NUMBER: US 60/196,267
? PRIOR FILING DATE: 2000-04-11
? PRIOR APPLICATION NUMBER: US 60/196,460
? PRIOR FILING DATE: 2000-04-11
? PRIOR APPLICATION NUMBER: US 60/196,527
? PRIOR FILING DATE: 2000-04-11
? PRIOR APPLICATION NUMBER: US 60/196,528
? PRIOR FILING DATE: 2000-04-11
? PRIOR APPLICATION NUMBER: US 09/687,837
? PRIOR FILING DATE: 2000-10-13
? PRIOR APPLICATION NUMBER: US 60/240,503
? PRIOR FILING DATE: 2000-10-13
? PRIOR APPLICATION NUMBER: US 60/240,508
? PRIOR FILING DATE: 2000-10-13
? PRIOR APPLICATION NUMBER: US 60/240,539
? PRIOR FILING DATE: 2000-10-13
? PRIOR APPLICATION NUMBER: US 60/240,543
? NUMBER OF SEQ ID NOS: 134
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1
? LENGTH: 7215
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: full length human CLASP-5 CDNA
? NAME/KEY: CDS
? LOCATION: (112)..(6159)
? OTHER INFORMATION: human CLASP-5
? US-09-736-960-1

Query Match          14.3%   Score 71.4; DB 10; Length 7215;
Best Local Similarity 51.1%; Pred. No. 6e-13;
Matches 208; Conservative 0; Mismatches 181; Indels 18; Gaps 1

QY      2 GTTTTACACCATCACCAAAACCAGATTCTTGATGAGATTAATAAATAGTTGCCACT 61
        ||| ||||| | | | | | ||||| || | ||||| || | ||||| ||
Db       1759 GTTACATTACCATTAATAGTCCTCCTGACTTTTAGAAGAAGTGAATAATTAACTCCCGCT 1818

QY      62 CAGGTGCATGAAAAGCAGCACCACTGTGGCTCACATCTTCCATGTCAGCTGTACAACCTCA 121
        |||| | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db       1819 AAGCTACAGTAGTAATCAACCACTCCTGTTGACACTTCTACCATATCAGCTGCAGCA---- 1874

QY      122 AGTAAAGAGAACGACGAGAGAGGGATGTCGTTGAAACCCAGTTGGCTACCTCGCTT 181
        ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db       1875 -----GAGGAGAGAGCTTCCTGGGAAACCTCCTGGGATATTCATGCGTG 1920

QY      182 CCCCTCTGTAAGAGAGGAGGGTGGTGACACAGACAGACATCCCGGTCTGGGAGAC 241
        || | || | | | | | | | | | | | | | | | | | | | | | |
Db       1921 CCAATTCTCTTAATAGAGAGCTTTCGAACCTGSAATCTTACTGTCTCCAGATTGCTTGGAA 1980

QY      242 CTTCCTCGGGCTATCTTGGCTACCAAGAGCTTGGAGTGGGAGGCATTAATGATCGGAA 301

```

```

Db      1981  A A A T T G C A C C C A C T A C T A C C A T T G C A T T G C T G C G A G A A A G T C C A C T T A A G A A T C C T C C C      2040
Qy      302  A T T A A T T G G T A G A T G A G G C A A G C C A C T G C T G A A A T T T C C A C T G A T C T G G T T C T A C A      361
Db      2041  A T T A A T G T G G C G G A A G A G A C A T A A G G A G A G A T T T A A T A T T G A A G T G C A A C T G T T C T T C T      2100
Qy      362  G T G T A T A C T C A G A T A C G A T T A C A T T A C A T A T T T T T C A G T C A T G T C A      408
Db      2101  G T A C A C A C C C A G A C A C A C C A C C T G G A G A A G T C T T A C C C T C T G C C A      2147

RESULT 4
US-09-736-968A-1
; Sequence 1, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-7 Transmembrane Protein
; FILE REFERENCE: 020054-000611US
; CURRENT APPLICATION NUMBER: US/09/736,968A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length human CLASP-7 cDNA
; NAME KEY: CDS
; LOCATION: (13)..(6156)
; OTHER INFORMATION: human CLASP-7
US-09-736-968A-1

Query Match          9.2%; Score 46; DB 93; Length 6372;
Best Local Similarity 63.6%; Pred. No. 0.00013;
Matches 70; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
Qy      2  G T T T A C A C C A T C A C C A A A C C A G A A T T T A T A T G A T T A A A T T A A A T T G A C T T G C C A C T      61

```

Db 1828 GTGGTCTACCAATCAAGTCCCGGAGTCTCTACGAGAGTTCAGACTGCATCTCCAGCC 1887
QY 62 CAGCTGCATGAAGACGACCGTGTGCACATCTCTCCATGTCAGCTG 111
Db 1888 TGGCTGACAGAGAACCATCACCCTGCTGTCCACCTTCTACCATGTGAGCTG 1937

RESULT 5

US-09-736-968A-77
; Sequence 77, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-7 Transmembrane Protein
; FILE REFERENCE: 020054-000611US
; CURRENT APPLICATION NUMBER: US/09/736,968A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: exon 88470-88680
US-09-736-968A-77

Query Match 8.1%; Score 40.4; DB 9; Length 211;
Best Local Similarity 65.6%; Pred. No. 0.0013;
Matches 59; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 22 CCCAGATTATGATGAGATTAAATAGATGCCCACTCAGCTGCATGAAGAAGCACA 81
Db 38 CCCAGATTATGAGAGAGATTCAAGCTTCATCCAGCTCGGTGACAGAGAACCATCA 97
QY 82 CCTGTGCTACATCTTCCATGTGACGCTG 111
Db 98 CCTGTGCTACCTTCTACCATGTGACGCTG 127

RESULT 6
US-09-070-927A-88
; Sequence 88, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven Barash

TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides

NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A

FILING DATE: 04-May-2000

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 88:

SEQUENCE CHARACTERISTICS:

LENGTH: 6691 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 88:

US-09-070-927A-88

Query Match 6.9%; Score 34.6; DB 10; Length 6691;
Best Local Similarity 54.3%; Pred. No. 0.78;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 12 ATCAACCAAAACCGAATTTATGATGAATTAATAGAGTTGCCCACTCAGCTGCATG 71
Db 1656 AACACCAAAAGAAATAGTCAAAAGAACTGATCAATATATTTGACAGCAAGCTGCCAA 1715
QY 72 AAAAGCACCACCTGTGTGTCACATTTCCATGTGACAGCTGTGACAACCTCAAGTAAAGAA 131
Db 1716 AAAATCAAGTGGCGGTAGCTTACGTAACCGCTATGCTGTGCAATTAAGAAATAT 1775
QY 132 GCACGAGA 140
Db 1776 GCAACAGA 1784

RESULT 7
US-09-738-626-795/C
; Sequence 795, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI


```

; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 795
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-795

Query Match
Best Local Similarity 54.1%; Score 32.4; DB 9; Length 1314;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
```

```

QY 40 GATTAAATAGAGTTGCCCTCAGTCGATGAAAGCACCACCTGTGCTCACAATCTT 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 GATTAGTAGAGTCTTTGACCGGACGACGATTAAGAGCGCTGTTTGACGGCCTT 833
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 CCATGTGACGTGTGACACCTCAAGTAAAGAGACGACGAAGAAGAGATGTCGTTGAAC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 CTTCACGGTGTGTCACCAAGTACCAACGAGGACCGGCTGGCTCGCGAC 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 160 CC 161
||
Db 772 CC 771
```

```

RESULT 8
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
```

```

; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match
Best Local Similarity 54.1%; Score 32.4; DB 9; Length 3309400;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
```

```

QY 40 GATTAAATAGAGTTGCCCTCAGTCGATGAAAGCACCACCTGTGCTCACAATCTT 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737638 GATTAGTAGAGTCTTTGACCGGACGACGATTAAGAGCGCTGTTTGACCGCCTT 737697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 100 CCATGTGACGTGTGACACCTCAAGTAAAGAGACGACGAAGAAGAGATGTCGTTGAAC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737698 CTTCACGGTGTGTCACCAAGTACCAACGAGGACCGGCTGGCTCGCGAC 737757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 160 CC 161
||
Db 737758 CC 737759
```

```

RESULT 9
US-09-771-208-20
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
```

```

; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; PRIOR FILING DATE: 1997-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (546998)..(547017)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (494715)..(494814)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (390986)..(391005)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (346600)..(346823)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (317174)..(317193)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (280353)..(280373)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (271829)..(271848)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (183872)..(183891)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
```

Query Match	6.3%;	Score 31.6;	DB 10;	Length 14141;
Best Local Similarity	56.9%;	Pred. No. 12;		

Db 295 TACTACACTCTGGATAATCATATAGATAATC

RESULT 12

US-09-938-842A-3143/C
; Sequence 3143, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: S000300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3143
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3143

Query Match

Best Local Similarity 6.2%; Score 31.2; DB 9; Length 2000;
Best Local Similarity 50.7%; Pred. No. 5.1;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 17 CAAACCCAGATTTTATGATGATTTAAATAGAGTTGCCACTGCGTCGATGAAG 76
DB 893 CAAAACCCAGATTTTATGATGATTTAAATAGAGTTGCCACTGCGTCGATGAAG 76
QY 77 CACACCTTCTCCTACATTTCTTCATGTCAGCTGTGACCACTCAAGTAAGAGCAG 136
DB 833 GGAAGATCTTGGGAGCGCTTTATCATGCGCAATTTAAGAGTAAGCAACCAATTTG 774
QY 137 AAGAAAGAGGATGCTGTTGAACCCAG 164
DB 773 AAGAGATTCCTCATTCATCAATCAAG 746

RESULT 13

US-09-960-352-12649/C
; Sequence 12649, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengding
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10288)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12649
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 54-LIB3058-010-Q1-K1-F6
US-09-960-352-12649

Query Match

Best Local Similarity 6.2%; Score 30.8; DB 10; Length 417;
Best Local Similarity 49.4%; Pred. No. 2.8;
Matches 80; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 323 AACCCACTGCTAAATTTCCATCTGCTTTCTACAGTGTATCTCAGATCAGCAT 382
DB 162 AACCTATAGTTAAATGAACAGTTTATATGAATGAAGCAGCATGATGTAAGGCTT 103

QY 383 TTACATTAATTTTTCAGTACTGTACAGAAACCGAATCTGGAGCCCAAGCTTAGGAAC 442

DB 102 TTTCATTAACATTTTTCAGTACTGTACAGAAACCGAATCTGGAGCCCAAGCTTAGGAAC 43

QY 443 GACCTGTAAGTCACTTAAGAGTCTGCATCGCATGGAAGC 484

DB 42 TAATCATTTAAATAAAGAGCGCGCTCGCATGTAGGAC 1

RESULT 14

US-09-764-864-1678
; Sequence 1678, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antihodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1678
; LENGTH: 32191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1678

Query Match

Best Local Similarity 6.1%; Score 30.6; DB 10; Length 32191;
Best Local Similarity 58.1%; Pred. No. 40;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 382 TTACATTAATTTTTCAGTACTGTACAGAAACCGAATCTGGAGCCCAAGCTTAGGAAC 441
DB 3777 TTACTATTTTGTGAAGGATTTTCATTAAGCTTAACATGTAGCTGAATTAATCTTAA 3836
QY 442 CGAAGTGTAAAGTCTTAAGAGTCTGATGC 474
DB 3837 TACATTAATTAATGATCAATTAATTTTGAATTC 3869

RESULT 15

US-09-939-980-228/C
; Sequence 228, Application US/09939980
; Patent No. US20020082234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenbery, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
; POLYPEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

: APPLICATION NUMBER: US/09/939,980
: FILING DATE: 27-Aug-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/936,165
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Gimmil, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P50549
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
: TELEX: <Unknown>
: INFORMATION FOR SEQ ID NO: 228:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1005 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
: SEQUENCE DESCRIPTION: SEQ ID NO: 228:
US-09-939-980-228

Query Match      6.1%; Score 30.4; DB 10; Length 1005;
Best Local Similarity 62.5%; Pred. No. 6.3;
Matches 65; Conservative 0; Mismatches 36; Indels 3; Gaps 1;

QY 2 GTTTACACCATCAGCAAAACCCGAAATTTATGA--TGAGATTAAATAGAGTTGGCC 58
DB 439 GTTGTCACACGACATATATACCTATATTTTCATGAGGTCAGATATGAAATGAGTACAC 380

-QY 59 ACTCAGCTGCATGAAGAAGCACCCCTGTTGCTCACAATTCCTTCCA 102
DB 379 ACTCCGCTGCATCTAAGACACCCCATCTTCCACTGTGACTGCA 336
```

Search completed: February 7, 2003, 09:14:04
Job time : 1286.29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 06:53:49 ; Search time 1008.32 Seconds
(Without alignments)
8030.908 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agtttaccatcaccaaa.....aggcacagtgatgcgcct 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estro:*
7: em_estro:*
8: em_hlc:*
9: qb_est1:*
10: qb_est2:*
11: qb_hlc:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estom:*
17: qb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	457.6	91.5	800	BE783911
2	359.8	72.0	465	BE938152
3	347.2	69.4	465	BE938144
4	308	61.6	891	BO919716
5	166	33.2	412	BB672508
6	166	33.2	441	BB819621

7	148.2	29.6	245	12	BF948041	MR3-NN021
8	144.4	28.9	538	12	BF472403	BF472403
9	134.8	27.0	543	10	BB763030	BB763030
10	117	23.4	345	10	BB222768	BB222768
11	106	21.2	330	10	BB207661	BB207661
12	100.2	20.0	285	10	BB225426	BB225426
13	94.4	18.9	274	10	BB177025	BB177025
14	90.4	18.1	314	10	BB219663	BB219663
15	89.8	18.0	249	10	BB538596	BB538596
16	83.6	16.7	212	10	AV344231	AV344231
17	80.6	16.1	255	10	BB200610	BB200610
18	78.2	15.6	259	10	BB241632	BB241632
19	77.4	15.5	219	10	BB184661	BB184661
20	76.6	15.3	917	17	CNS03GUC	CNS03GUC
21	74	14.8	447	12	BG019353	BG019353
22	73.2	14.6	239	10	BB256604	BB256604
23	57	11.4	669	13	BI961883	BI961883
24	55.6	11.1	749	13	BI828723	BI828723
25	53.4	10.7	147	14	BQ339280	BQ339280
26	51.6	10.3	271	13	BM029853	BM029853
27	49.2	9.8	534	12	BE945462	BE945462
28	49.2	9.8	656	13	BQ395781	BQ395781
29	46	9.2	909	14	BQ730710	BQ730710
30	45.6	9.1	695	14	C88441	C88441
31	42.4	8.5	254	17	AA391952	AA391952
32	40.8	8.2	657	13	AE632934	AE632934
33	39	7.8	1507	11	AK010755	AK010755
34	39	7.8	522	10	BE667330	BE667330
35	37.6	7.5	631	13	BI507332	BI507332
36	37.6	7.4	635	9	AA892774	AA892774
37	37	7.4	642	14	BO206156	BO206156
38	37	7.4	753	14	BO201592	BO201592
39	37	7.4	836	17	CNS0090K	CNS0090K
40	36.6	7.3	802	17	CNS01781	CNS01781
41	36.2	7.2	350	10	AW784142	AW784142
42	36	7.2	595	9	AA695066	AA695066
43	36	7.2	864	13	BI558960	BI558960
44	36	7.2	864	13	BI558960	BI558960
45	35.4	7.1	995	17	CNS02102	CNS02102

ALIGNMENTS

RESULT 1
LOCUS BE783911 800 bp mRNA linear EST 20-OCT-2000
DEFINITION 601470964F1 NIH_MGC_67 Homo sapiens cdna clone IMAGE:3874046 5',
ACCESSION BE783911 mRNA sequence.
VERSION BE783911.1 GI:10205109
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 800)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fmail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LAM9630 row: p column: 15
High quality sequence stop: 629.
Location/Qualifiers 1..800

FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3874016"
/clone_id="NH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: oligo dfr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT      227 a      198 c      185 g      190 t
ORIGIN

Query Match      91.5%; Score 457.6; DB 12; Length 800;
Best Local Similarity 98.6%; Pred. No. 6.9e-133;
Matches 493; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

4 TTATACCATCTACCAAAACCCAGATTATTTATGATGAGATTAAATAGAGTTGCCACTC 62
|||||
20 TTATACCATCTACCAAAACCCAGATTATTTATGATGAGATTAAAGCTAGAGTTGCCACTC 79
|||||
63 AGCTGATGAAAGACCAACCTGTTGCTACATTTCTCCATGTCAGCTGTGACAAC-TCA 121
|||||
80 AGCTGATGAAAGACCAACCTGTTGCTACATTTCTCCATGTCAGCTGTGACAACGTCA 139
|||||
122 AGTAAAGAGACGAGAGAGAGGATGCTGTGAACCCAGTTGGCTACGCTGGCTT 181
|||||
140 AGTAAAGAGACGAGAGAGAGGATGCTGTGAACCCAGTTGGCTACGCTGGCTT 199
|||||
182 CCCCTCTGAAAGACGAGAGGATGCTGTGAACGAGACACATCCGGTCTGGCGAAC 241
|||||
200 CCCCTCTGAAAGACGAGAGGATGCTGTGAACGAGACACATCCGGTCTGGCGAAC 259
|||||
242 CTTCCTTGGGCTATCTTGGCTACCAAGCTTGGGATGGGACGATTAATGGTCCGAA 301
|||||
260 CTTCCTTGGGCTATCTTGGCTACCAAGCTTGGGATGGGACGATTAATGGTCCGAA 319
|||||
302 ATTAATGGGTAGATGAGAGCAAGCCACTGTGAA-ATTTCCACTCATCTGGTTCTAC 360
|||||
320 ATTAATGGGTAGATGAGAGCAAGCCACTGTGAAATTTCCACTCATCTGGTTCTAC 379
|||||
361 AGTGTATCTAGATGAGATGATTTACATATTTTTCAGTACTGTGAGAAACCGAATC 420
|||||
380 AGTGTATCTAGATGAGATGATTTACATATTTTTCAGTACTGTGAGAAACCGAATC 439
|||||
421 TGGAGCCCAAGGCTTGGAGAGCAACTGTAAAGTACCTTAAGATGAGTGGCATGGA 480
|||||
440 TGGAGCCCAAGGCTTGGAGAGCAACTGTAAAGTACCTTAAGATGAGTGGCATGGA 499
|||||
481 AGGCCAGCTGATGATGCCT 500
|||||
500 AGGCCAGCTGATGATGCCT 519
|||||

RESULT 2
BE938152      465 bp      mRNA      linear      EST 02-OCT-2000
LOCUS      MR3-TN0048-280800-001-g08 TN0048 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BE938152
ACCESSION      BE938152.1 GI:10465154
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 465)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
```

```
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
MEDLINE
20202663
CONTACT: Simpson A.J.G.
COMMENT      Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=tl2-MR3-TN0048-280
800-001-g08&cl=2000-08-28&cl=1)
Seq primer: puc 18 forward
High quality sequence stop: 465.

FEATURES
source
1..465
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="TN0048"
/dev_stage="Adult"
/notes="Organ: testis; normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      127 a      114 c      115 g      109 t
ORIGIN

Query Match      72.0%; Score 359.8; DB 12; Length 465;
Best Local Similarity 99.4%; Pred. No. 3.5e-102;
Matches 361; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AGTTTACACCATCACCAAAACCCAGATTATTTATGATGAGATTAAATAGAGTTGCCAC 60
|||||
103 AGTTTACACCATCACCAAAACCCAGATTATTTATGATGAGATTAAATAGAGTTGCCAC 162
|||||
61 TCAGCTGATGAAAGACCAACCTGTTGCTACATTTCTCCATGTCAGCTGTGACAATC 120
|||||
163 TCAGCTGATGAAAGACCAACCTGTTGCTACATTTCTCCATGTCAGCTGTGACAATC 222
|||||
121 AAGTAAAGAGACGAGAGAGAGGATGCTGTGAACCCAGTTGGTACTCTGGCT 180
|||||
223 AAGTAAAGAGACGAGAGAGAGGATGCTGTGAACCCAGTTGGTACTCTGGCT 282
|||||
181 TCCCTCTCTGAAAGACGAGAGGATGCTGTGAACGAGACACATCCGGTCTGGCGAA 240
|||||
283 TCCCTCTCTGAAAGACGAGAGGATGCTGTGAACGAGACACATCCGGTCTGGCGAA 342
|||||
241 CCTTCCTTGGGCTATCTTGGCTACCAAGCTTGGATGGGACGAGATTAATGGTCCGA 300
|||||
343 CCTTCCTTGGGCTATCTTGGCTACCAAGCTTGGATGGGACGAGATTAATGGTCCGA 402
|||||
301 AATTAATGGGTAGATGAGAGCAAGCACTGCTGAATTTTCACATCTGTTCTAC 360
|||||
403 AATTAATGGGTAGATGAGAGCAAGCACTGCTGAATTTTCACATCTGTTCTAC 462
|||||
361 AGT 363
|||||
463 AGT 465
|||||

RESULT 3
BE938144      465 bp      mRNA      linear      EST 02-OCT-2000
LOCUS      MR3-TN0048-280800-001-e05 TN0048 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BE938144
ACCESSION      BE938144.1 GI:10465135
VERSION
KEYWORDS
```

Source	Organism	Human
REFERENCE	Authors	Human
1	(bases 1 to 465)	
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.V. and Simpson, A.J.G.		
Shogun sequencing of the human transcriptome with ORF expressed sequence tags		
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
20202663		
CONTACT: Simpson A.J.G.		
Laboratory of Cancer Genetics		
Ludwig Institute for Cancer Research		
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil		
Tel: +55-11-2704922		
Fax: +55-11-2707001		
Email: asimpson@ludwig.org.br		
This sequence was derived from the FAESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL		
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-WR3-TN0048-280800-001-e056t3=2000-08-28&t=1)		
Seq primer: puc 18 forward		
High quality sequence stop: 465.		
Location/Qualifiers		
1..465		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/clone_lib="TN0048"		
/dev_stage="Adult"		
/note="Organ: testis_normal; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - ludwig institute for Cancer research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
BASE COUNT	129 a	112 c 116 g 108 t
ORIGIN		
Query Match	69.4%	Score 347.2; DB 12; Length 465;
Best Local Similarity	98.9%	Pred. No. 3.3e-98;
Matches 360; Conservative	0; Mismatches	3; Indels 1; Gaps 1;
QY 1	AGTTTACACACATCACCAAAACCCAGATTTTATGATGAGATTAAATAAGAGTTGCCAC	60
Db 102	AGTTTACACACATCACCAAAACCCAGATTTTATGATGAGATTAAATAAGAGTTGCCAC	161
Y 61	TTCAGCGCANTGAAAGAGCACCACCTTCTGTCACATCTTCCATGTCAGCTGACACATC	120
Db 162	TTCAGCTGCATGAAAGACACACCTTCTGTCACATCTTCCATGTCAGCTGACACATC	221
QY 121	AAGTAAAGACACGACGAAAGAGGGATGTCTTGAACCCAGTTGGCTACTCTGCT	180
Db 222	AAGTAAAGACGACGAAAGAGGGATGTGATGAAACCCAGTTGGCTACTCTGCT	281
QY 181	TCCGCTTCGTAAGAGCGAAGGG - TGTGTACAAACGACGACACATCCGGGTCTGGGCA	239
Db 282	TCCGCTTCGTAAGAGCGAAGGGTGTGTGAAACGACGACACATCCGGGTCTGGGCA	341
QY 240	ACCTTCCTCGGGCTATCTTGGCTACCAAGAGCTTGGAGTGGGACAGCATTTATGTCGG	299
Db 342	ACCTTCCTCGGGCTATCTTGGCTACCGAGAGCTTGGAGTGGGACAGCATTTATGTCGG	401
QY 300	AAATTAATGAGTATGAGGCAACGACATCTCTGAAAAATTTCCATCATCTGTGTTCTA	359
Db 402	AAATTAATGAGTATGAGGCAAGCCACATCTCTGAAAAATTTCCATCATCTGTGTTCTA	461

XY	360	CAGT	363	
Db	462	CAGT	465	
RESULT 4				
LOCUS	B0919716			
DEFINITION	B0919716	891 bp	mRNA	linear
ACCESSION	AGENCOURT_8850966	NCI_CGAP_Man2	Mus musculus	cdna clone
VERSION	B0919716			
KEYWORDS	B0919716.1	GI:22334414		
SOURCE	EST.			
ORGANISM	house mouse.			
REFERENCE	Mus musculus			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
JOURNAL	1 (bases 1 to 891)			
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .			
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strusberg, Ph.D.			
	Email: cgapbs-remail.nih.gov			
	Tissue Procurement: Gilbert Smith, Ph.D.			
	CDNA Library Preparation: Life Technologies, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNL at:			
	http://image.llnl.gov			
	Plate: LLAM13965 row: 0 column: 23			
	High quality sequence stop: 412.			
FEATURES				
source	location/Qualifiers			
	1..891			
	/organism="Mus musculus"			
	/strain="FVB/N-3"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:6441934"			
	/clone_lib="NCI_CGAP_Man2"			
	/tissue_type="tumor, biopsy sample"			
	/dev_stage="5 months"			
	/lab_host="DH10B"			
	/note="Organ: mammary; Vector: pCMV-SPOrt6; Site:1: SalI;			
	Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.			
	Library constructed by Life Technologies. Investigator			
	providing samples: Gilbert Smith, NIH"			
BASE COUNT	243 a	209 c	236 g	201 t
ORIGIN				
Query Match	61.6%;	Score 308;	DB 14;	Length 891;
Best Local Similarity	87.1%;	Pred. No. 1e-85;		
Matches	338;	Conservative	0;	Mismatches 50; Indels 0; Gaps 0;
113	GACAACTCAAGTAAAGAGACGACGAAGAAGAGGATGTCTGTTGAACCCAGATTGGCTAC	172		
1	GATATCTCCACCAAGAGAGACGAGAAAGAGAGCGCTGTGGAAAGCGAGGTGGCTTT	60		
173	TCTGGGCTTCCCTCCCTGAGAAAGCGAAGGGTGTGGAACAACGAGCAGACATCCCGCTC	232		
61	TCTCGGTGCTCTCTCTGAAAGATGGAAGGGTGTGAAGATGAGACGACATCCCGCTC	120		
233	TGCGGGAACCTTCCTTCGCGGCTATCTTGGCTACCAAGAGCTTGGGATGGCGAGCATTA	292		
121	TGCGCTAACCTGCCATCTGGCTACCTCGGCTACCGAGAGCTCGGCATGGCGAGCATTA	180		
293	GGTCGGGAATTTAAATGGGTAGTGGAGGCAAGCCACGTCGAAATTTCCACATCATTC	352		
181	GGTCAAGAGTTAAAGTGGGTGGAAGAGGACAGCCACGCTGTAAGATCTCCACTCATCTG	240		
353	GTATTCTACAGTGTATCTCAGATCAGCATTCATTAATTTTTCGAGTACTGTCAAAA	412		
241	GTATTCACAGTGTATCTCAGATCAGCATTCATTAATTTTTCGAGTACTGTCAAAA	300		

BASE COUNT	130 a	100 c	76 g	106 t
ORIGIN				
Query Match	33.2%; Score 166; DB 10; Length 412;			
Best Local Similarity	63.1%; Pred. No. 3.8e-41;			
Matches 256; Conservative	0; Mismatches 150; Indels 0; Gaps 0;			
Oy	2 GTTTTACCATCACCACCAACCCAGATTTTATGATGATGATTAATAGAGTGGCCACT 61			
Db	7 GTGTACACCATTTCTCAAAATCCACATTTCTCTATGATGAGTGAATTAATGACATCCACAA 66			
Oy	62 CAGCTGCATGAAGAAGCACCCTGTGTCTCACAATTTTCCATGTGCAGCTGTGACAATCA 121			
Db	67 CAACTCCATGGAAGAACATCACCCTTTGTTCTCTTTTACACATCAATGATGACATCAAT 126			
Oy	122 AGTAAAGGAAGCAGGAAGAGGAGATGTCTGTAAACCCAGTTGGCTACTCTGGCTT 181			
Db	127 GCCAAAGCCATGCCAAAGAAAGAGCTTTGGAGCATAGGGAGCTATGATGATGCTT 186			
Oy	182 CCCCCTCTGAAACAGCGAAGSGTGTGTACAAAGCAGACACATCCCGCTCTGGCGAAC 241			
Db	187 CCTCTGATGAAACATATATCAATTAATCTTCTCAGAGGTACACATCCCAATAGACAGGACC 246			
Oy	242 CTTCTCTGGGCGCTTTTGTGCTACCAAGAGCTTGGAGTGGCAGGACATTAATGTCCGGA 301			
Db	247 CTGCTCTCAATTAATTTTAAGCATTCACATGCTCTCAAGTGAAGCATGTGTGAATGTAC 306			
Oy	302 ATTAAATGGGTAAATGAGGACGACGACGTCGMAAATTTCCACTCATCTGGTTCTTACA 361			
Db	307 ATTAAATGGGTGATGTGGCAACCGCTTTTCAAGTGTTCACATTTGTGTGATCAACA 366			
Oy	362 GTGTATACCTCAGATCAGCATTTTACATTAATTTTTTCCAGTACTGTC 407			
Db	367 GTGAACACTCAGACCCCATGTAATGCATTTTCCGTAGTGGC 412			
RESULT 6				
LOCUS	BB819621 441 bp mRNA linear EST 19-NOV-2001			
DEFINITION	BB819621 Riken full-length enriched, lung RCB-0558 LLC CDNA Mus			
ACCESSION	musculus CDNA clone G730050N08 3', mRNA sequence.			
VERSION	BB819621.1 GI:16992250			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 441)			
	Akimura,T., Aikawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayasato,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,Y., Kojima,Y., Kondo,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahita,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yatsunishi,A., Yumetani,M. and Hayashizaki,Y.			

Best Local Similarity 89.9%; Pred. No. 1.2e-35;
Matches 204; Conservative 0; Mismatches 18; Indels 5; Gaps 4;

QY 195 ACGAAGGGTGTGACAA-GCGAGCAGACATCCGCTCGCGGACCTTCCTCGGCG 253
DB 19 ACGAAGGGTGTGACAAAGCGAGCATATCCCGCTCGCGGACCTTCCTCGGCG 78
QY 254 TATCT-TGCGTACCAAGAGCTTGAGTGGGAGCATTTAGTCCGGAATTAATGGCT 312
DB 79 GATCTGTGATACAGAGAGCTTGAGTGGGAGCATTTAGTCCGGAATTAATGGCT 138
QY 313 AGATGAGAGCAGACCTGCTGAAATTTTC-CACATATCTGCTTTTACAGGTATAC 371
DB 139 AGATGAGAGCAGACCTGCTGAAATTTTC-CACATATCTGCTTTTACAGGTATAC 198
QY 372 AGGATCAGCATTTTACATATTTT--TTCCAGTACTGTCAGAAACCG 416
DB 199 AGGATCAGCATTTTACATATTTTGTCTCATGTCATGTCAGAAACCG 245

RESULT 8
LOCUS BF472403 538 bp mRNA linear EST 04-DEC-2000
DEFINITION UT-M-BH3-awc-c-08-0-UT.1 NIH_BMAP_M.S4 Mus musculus cDNA clone
ACCESSION BF472403
VERSION BF472403.1 GI:11541586
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 538)
TITLE Bernaldo,M.F., Lennon,G. and Soares,M.B.
JOURNAL Normalization and subtraction: two approaches to facilitate gene
MEDLINE discovery
COMMENT Genome Res. 6 (9), 791-806 (1996)
9704477

CONTACT: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890

Email: MEST@mail.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
Source 1..538

Organization="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UT-M-BH3-awc-c-08-0-UT"
/clone.lib="NIH_BMAP_M.S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; The
NIH_BMAP_M.S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M.S4,
NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,

NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
(NIH_BMAP_M.S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M.S4 library. This procedure has been previously
described (Bernaldo, Lennon and Soares, Genome Research
6:791-806, 1996)"

BASE COUNT 150 a 122 c 123 g 143 t
ORIGIN
Query Match 28.9%; Score 144.4; DB 12; Length 538;
Best Local Similarity 90.6%; Pred. No. 2.8e-34;
Matches 154; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 331 GCTGAATTTTCCACTATCTGTTCTACAGTGTATCTACGAGATTCATTAATTA 390
DB 13 GTTGAAGATCTCCACTATCTGTTCTACAGTGTATCTACGAGATTCATTAATTA 72
QY 391 TTTTTCAGTACTGTGAGAAACCGAATCTGGAGCCCAAGCCTTAGAAGCACTGCT 450
DB 73 TTTTTCAGTACTGTGAGAAACCGAATCTGGAGCCCAAGCCTTAGAAGCACTGCT 132
QY 451 AAGTACTCTTAAGAGTCTGCATGCGATGAGAGGCCACGATGATATGCGCT 500
DB 133 AAATATCTTAAGAGTCTGCATGCGATGAGAGGCCACGATGATATGCGCT 182

RESULT 9
LOCUS BF763030 543 bp mRNA linear EST 17-OCT-2001
DEFINITION BF763030 RIKEN full-length enriched, B16 F10Y cells Mus musculus
ACCESSION BF763030
VERSION BF763030.1 GI:16207943
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 543)
TITLE Akimura,T., Arakawa,T., Carinici,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Matahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

Unpublished (2001)
Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome-res@sc.riken.go.jp/
Carinici,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

FEATURES

Location/qualifiers

BASE COUNT	ORIGIN
164 a	137 c
	101 g
	141 t

Query Match	27.0%;	Score 134.8;	DB 10;	Length 543;
Best Local Similarity	63.0%;	Pred. No. 3e-31;		
Matches 257; Conservative	0;	Mismatches 147;	Indels 4;	Gaps 3;

QY	2	GTATTACCACTACCAAAACCCGAGATTTTATGATGAGATTAATAATAGTGGCCACT	61
Db	138	GTGTGACACACTTCTCCAAATTCAGATTTCTGTATGAGGTAAATATGAACTACCAACA	197
QY	62	CACCTGCATGAAAAGACACACCTGTTCTCATATCTTGCATGTCAGCTGTACACTCA	121
Db	198	CACCTGCATGAAA -CATCACCTTTTGTCTCTTTACACATCATCATGTGACATCAAT	256
QY	122	AGTAAAGSACGACAGAGAAGGAGTGTCTTGAATCCCAAGTTGGCTACTCTGGCTT	181
Db	257	GCCAAAGCCAAATGCCAAAAGAGAGCTTTGGAACATCATGTGGCTATGACATGGCTT	316
QY	182	CCCCCTCCTAAAGACGGAAGGGTGTACAAAGCAGCAGCATATCCCGTCTCGGCGAC	244
Db	317	CCCTCTATGAAAACATGATCAATATAGCTTTCAGAGATATCAACATCCCAATACCAACCC	376
QY	242	CTTCTCTCGGGCTATCTTGGCTACCAAGACCTTGGAGTGGGCG -AGGATTAATGGTCGG	299
Db	377	CTCCCTCCCTAATTTATTTAAGCATTCAGATTCCTAGCAAGTCCAAAGGCATGGTGAAGTG	439
QY	300	AAATTAATGGGTAGATGAGGCAAGCCACTGCTGAAAAATTCACATCATCTGGTTCTA	359
Db	437	ACATTAATAATGGTTGCAT -GTGGCAAAACCGGCTTTCAAAGTGCACAAATTTGTGTATCAA	495
QY	360	CATGTGATACTCAGAGTCAGCATTTAATATATTTTTTCCAGTACGTGC 407	
Db	496	CATGTGAACATCAGGACCCCAATGATTAATTCATTTTTTCCGATAGGCC 543	

RESULT 10	BB222768			
LOCUS	BB222768			
DEFINITION	BB222768	345 bp	mRNA	linear
ACCESSION	musculus cDNA clone A530079G09 3', mRNA sequence.			EST 01-JUL-2000
VERSION	BB222768			
KEYWORDS	BB222768.1	GI:8891380		
EST				

SOURCE ORGANISM	REFERENCE
house mouse, <i>Mus musculus</i>	1 (bases 1 to 345)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.	Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arahawa, T., Carninci
AUTHORS	

TITLE	RIKEN Mouse ESTs (Konno, H., et al.)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC) Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-72 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, Y., N. Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermotranslabilization and thermooxidation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)
Itoh, M., Kitsunuma, T., Akiyama, J., Shibata, K., Izawa, N., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-phased high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)
Please visit our web site (<http://genome.ritc.riken.go.jp/>) for further details.

FEATURES	Location/Qualifiers
Source	1. .345

BASE COUNT	105 a	92 c	61 g	87 t
ORIGIN				

VERSION	BB225426.1	GI:8894037
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 285)	
AUTHORS	Konno,H., Aizawa,K., Akahira,S., Akiyama,T., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Ishikawa,T., Itoh,M., Izawa,M., Kadote,K., Kagawa,I., Kai,C., Kawai,T., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,I., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shibemoto,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomiなが,N., Toya,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,T., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.	
TITLE	RIKEN Mouse ESTs (Konno,H., et al.)	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: yoshihide Hayashizaki	

BB225426.1 GI:8894037
EST.
house mouse.
Mus musculus
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata: Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 285)
Kono H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci
P., Endo, T., Fukuda, S., Fukunishi, Y., Hata, A., Hayatsu, N.,
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, Y., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata
Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya
T., Tsunoda, Y., Watanishi, A., Watanabe, S., Yamamura, T., Yamanaka, I.,
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-2 Tsukuba-cho, Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermosensitization and thermoactivation of thermolabile enzymes by
cDNA libraries and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details

```

Location/Qualifiers
1. 285
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A53092E21"
/clone_lib="RIKEN full-length enriched, adult male aorta
and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory
in RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAGAGATCCACAGAGCTTTTTTTTTTTTTTNN 3']. cDNA was
prepared by using retrolose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 439.0. Second
strand cDNA was prepared with the primer adaptor of
sequence [5' GAGGAGAGAGTTCTCGATTAATTAATTAATATCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pbluescript KS(+) after bulk excision from Lambda
UDC I."

```

BASE COUNT	88 a	68 c	60 g	69 t
ORIGIN				
Query Match	20.0%; Score 100.2; DB 10; Length 285;			
Best Local Similarity	59.8%; Pred. No. 1.8e-20;			
Matches 168; Conservative	0; Mismatches 113; Indels 0; Gaps			
QY	125	AAAGGAGCAGCAGAGAAGAGGAGGATGCTCGTTGAAACCCAGTTGGCTACTCTCGGCTTCCC	184	
Db	4	AAAGCCCATGCCAAATTAAGAAGAGGCTTCGGAGACATATAGTGGGCTATGCATGGCTTCAT	63	
QY	185	CTCTCTGAAGAGCGAAGGGTGGTGCAAGCAGCAGCATCCCGCTCTCGGCGAACCTT	244	
Db	64	CTGTGAAAACATGATCAATATAGCTTCTCAGAGATGACAAATCCCAATAGCAACGACCTG	122	
QY	245	CCTTGGGGCTATCTTGGCTACCAAGAGCTTGGGATGGCGAGCATTTATGTGCGGAAATT	304	
Db	124	CTCCCAATTAATTTAAGCATTTCAAGATCTCTCAAGGTGCAAGACATGGTGGGAAGGACATT	183	
QY	305	AAATGGGTAGATGGAGGCGAAGCGCACTGCTGAAATTTTCCATCTCATCTGGTGTTCACAGTG	364	
Db	184	AAATGGGTGATGGTGGGCAAAACCGCTTTTCAAGTGTCCACATTTGTGTATCAACAGTG	243	
QY	365	TATACTCAGGATCAGCATTTTCATATATTTTTTCCAGTACTG	405	
Db	244	AACACTCAGGACCCACATGTAAATGCATTTTCGTGTCAGTG	284	

RESULT 13	
BBI77025	274 bp mRNA linear EST 29-JUN-2000
LOCUS	
DEFINITION	BBI77025 RIKEN full-length enriched, adult male hypothalamus Mus musculus CDNA clone A230064N07 3' , mRNA sequence.
ACCESSION	m8720ac

VERSION	BB1/7023.1
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Garninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the syntheses of full length cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)
Itoh, M., Ritsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Garninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:39:15 ; Search time 996.663 Seconds

(without alignments)
14629.322 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000

Perfect score: 501

Sequence: 1 ttacactgaaagaagtcct.....gaggaagccctccatgatgga 501

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	499.4	99.7	2469	AK090793	AK090793 Homo sapi
2	499.4	99.7	3899	AX058220	AX058220 Sequence
3	475.4	94.9	2516	AB049841	AB049841 Macaca fa
4	426.4	85.1	7506	AX255048	AX255048 Sequence
5	426.4	85.1	7522	AF527605	AF527605 Homo sapi
6	426.4	85.1	7545	AB028981	AB028981 Homo sapi
7	366.2	73.1	2998	BC009134	BC009134 Mus muscu
8	359.8	71.8	3227	RNMR	X68101 R.norvegicu
9	274.4	54.8	2610	AX058192	AX058192 Sequence
10	274.4	54.8	3152	AK056684	AK056684 Homo sapi
11	274.4	54.8	6454	AX173022	AX173022 Sequence
12	269.6	53.8	3472	AX173120	AX173120 Sequence
13	269.6	53.8	4391	AX173028	AX173028 Sequence
14	269.6	53.8	4393	AX173118	AX173118 Sequence
15	263.4	52.6	2299	AK054649	AK054649 Homo sapi
16	248	49.5	13987	CNS01RGX	AL160233 Human chr
17	243.4	48.6	2768	AB056820	AB056820 Macaca fa
18	222.2	44.4	16316	AL161420	AL161420 Human DNA
19	216.4	43.2	2413	BC015018	BC015018 Homo sapi
20	216.4	43.2	2422	AK000227	AK000227 Homo sapi
21	190.6	38.0	12825	AC119357	AC119357 Rattus no
22	179.4	35.8	182640	AC126253	AC126253 Mus muscu
23	171.4	34.2	10103	AC020032	AC020032 Drosophila
24	171.4	34.2	159350	AC008354	AC008354 Drosophila
25	171.4	34.2	257224	AE003617	AE003617 Drosophila
26	168.2	33.6	1476	AX399260	AX399260 Sequence
27	146.2	29.2	153026	AL391280	AL391280 Human DNA
28	143	28.5	147556	AC110359	AC110359 Rattus no
29	140.2	28.0	162378	AC105541	AC105541 Rattus no
30	138.6	27.7	168684	AC023985	AC023985 Homo sapi
31	138.6	27.7	171811	AC011739	AC011739 Homo sapi
32	138.2	27.6	181127	AL672038	AL672038 Mouse DNA
33	131.6	26.3	174041	AL772340	AL772340 Danio rer
34	116.4	23.2	2149	AX173233	AX173233 Sequence
35	116.4	23.2	2825	BC008335	BC008335 Homo sapi
36	116.4	23.2	6372	AX173175	AX173175 Sequence
37	112.8	22.5	1900	BC021414	BC021414 Mus muscu
38	102.6	20.5	4886	AB037816	AB037816 Homo sapi
39	93.4	18.6	2303	BC024917	BC024917 Mus muscu
40	93.4	18.6	2972	BC024823	BC024823 Mus muscu
41	88.2	17.6	73515	AC017375	AC017375 Drosophila
42	88.2	17.6	166626	AC008318	AC008318 Drosophila
43	88.2	17.6	301639	AE003590	AE003590 Drosophila
44	83.2	16.6	2014	AX174662	AX174662 Sequence
45	83.2	16.6	4026	AX174571	AX174571 Sequence

ALIGNMENTS

RESULT 1
AK090793
LOCUS
DEFINITION
AK090793 Homo sapiens cDNA FLJ33474 fis, clone BRAMY200269, highly similar
to Trg protein.
ACCESSION
AK090793.1 GI:21749019
VERSION
Oligo capping; fis (full insert sequence).
KEYWORDS
Homo sapiens amygdala cDNA to mRNA, clone_11b:BRAMY2
SOURCE
clone:BRAMY2002369.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1
Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,

Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Kanda,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kawaguchi,A., Wagahtsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2469)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUN-2002) Takao Isogai, FLJ Project(HRI Team): 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomeshri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-6 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2469
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRAM2002369"
/issue_type="amygdala"
/clone_lib="BRAMy2"
/note="Cloning vector: PME18SFL3"
BASE COUNT 666 a 561 c 556 g 686 t
ORIGIN
Query Match 99.7%; Score 499.4; DB 9; Length 2469;
Best Local Similarity 99.8%; Pred. No. 2.7e-122;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTACACTGGAAGAAGTCTTTGTCGGACACATTGCAAGCATCATCTGTGAGCA 60
DB 1641 TTACACTGGAAGAAGTCTTTGTCGGACACATTGCAAGCATCATCTGTGAGCA 1700
QY 61 GCTGATAGCAGACGTTGTGGCATTGGGAAACAGATTCCAGAGTCCCTGTCCATCAT 120
DB 1701 GCTGATAGCAGACGTTGTGGCATTGGGAAACAGATTCCAGAGTCCCTGTCCATCAT 1760
QY 121 CAACAACCTGTGCCAACAGTACGACGCGCTTATTAAAGCACACACAGCTTCTCTGATGTGAA 180
DB 1761 CAACAACCTGTGCCAACAGTACGACGCGCTTATTAAAGCACACACAGCTTCTCTGATGTGAA 1820
QY 181 GGACTTAACCAAAAGATCGCAGCGTGTATGGCCACCGCCGAGTAAGAGAGCATGA 240
DB 1821 GGACTTAACCAAAAGATCGCAGCGTGTATGGCCACCGCCGAGTAAGAGAGCATGA 1880
QY 241 GAAGCACCAGAGATGCTGGTGAACCTCCAGTACAGCCTGGCCAAATCCTATGCGAGCA 300
DB 1881 GAAGCACCAGAGATGCTGGTGAACCTCCAGTACAGCCTGGCCAAATCCTATGCGAGCA 1940
QY 301 GCCCGAGCTCAGGAAGAGCTGGCTGACAGCATGGCCAGAGATCCATGTCAAAAATGGCGA 360
DB 1941 GCCCGAGCTCAGGAAGAGCTGGCTGACAGCATGGCCAGAGATCCATGTCAAAAATGGCGA 2000
QY 361 TCTCTAGAGGAGCAGATGCTATGTCCAGTACAGCCCTAGTGGAGATATCTCAC 420
DB 2001 TCTCTAGAGGAGCAGATGCTATGTCCAGTACAGCCCTAGTGGAGATATCTCAC 2060
QY 421 ACGGAAAGGCGTGTATTAGACAAGAGATGCAACGCTTTCAGGGTCATTACCCCAACATGCA 480
DB 2061 ACGGAAAGGCGTGTATTAGACAAGAGATGCAACGCTTTCAGGGTCATTACCCCAACATGCA 2120
QY 481 CGAGGAGCGCTCCATGATGCA 501
DB 2121 CGAGGAGCGCTCCATGATGCA 2141

RESULT 2
AX058220 3899 bp DNA linear PAT 17-JAN-2001
LOCUS
AX058220 Sequence 90 from Patent WO0077040.
DEFINITION
AX058220
ACCESSION
AX058220.1 GI:12310721
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 3899)
AUTHORS
Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Banduhn,O., Baughn,M.R., Azimzai,Y., Yang,D., Reddy,R. and Lu,D.A.
TITLE
Human intracellular signaling molecules
JOURNAL
Patent: WO 0077040-A 90 21-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..3899
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1078 a 929 c 949 g 943 t
ORIGIN
Query Match 99.7%; Score 499.4; DB 6; Length 3899;
Best Local Similarity 99.8%; Pred. No. 2.7e-122;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTACACTGGAAGAAGTCTTTGTCGGACACATTGCAAGTCATCATCTGTGAGCA 60
DB 1922 TTACACTGGAAGAAGTCTTTGTCGGACACATTGCAAGTCATCATCTGTGAGCA 1981
QY 61 GCTGATAGCAGACGTTGTGGCATTGGGAAACAGATTCCAGAGTCCCTGTCCATCAT 120
DB 1982 GCTGATAGCAGACGTTGTGGCATTGGGAAACAGATTCCAGAGTCCCTGTCCATCAT 2041
QY 121 CAACAACCTGTGCCAACAGTACGACGCGCTTATTAAAGCACACACAGCTTCTCTGATGTGAA 180
DB 2042 CAACAACCTGTGCCAACAGTACGACGCGCTTATTAAAGCACACACAGCTTCTCTGATGTGAA 2101
QY 181 GGACTTAACCAAAAGATCGCAGCGTGTATGGCCACCGCCGAGTAAGAGAGCATGA 240
DB 2102 GGACTTAACCAAAAGATCGCAGCGTGTATGGCCACCGCCGAGTAAGAGAGCATGA 2161
QY 241 GAAGCACCAGAGATGCTGGTGAACCTCCAGTACAGCCTGGCCAAATCCTATGCGAGCA 300
DB 2162 GAAGCACCAGAGATGCTGGTGAACCTCCAGTACAGCCTGGCCAAATCCTATGCGAGCA 2221
QY 301 GCCCGAGCTCAGGAAGAGCTGGCTGACAGCATGGCCAGAGATCCATGTCAAAAATGGCGA 360
DB 2222 GCCCGAGCTCAGGAAGAGCTGGCTGACAGCATGGCCAGAGATCCATGTCAAAAATGGCGA 2281
QY 361 TCTCTAGAGGAGCAGATGCTATGTCCAGTACAGCCTTTCAGGGTCATTACCCCAACATGCA 420
DB 2282 TCTCTAGAGGAGCAGATGCTATGTCCAGTACAGCCTTTCAGGGTCATTACCCCAACATGCA 2341
QY 421 ACGGAAAGGCGTGTATTAGACAAGAGATGCAACGCTTTCAGGGTCATTACCCCAACATGCA 480
DB 2342 ACGGAAAGGCGTGTATTAGACAAGAGATGCAACGCTTTCAGGGTCATTACCCCAACATGCA 2401
QY 481 CGAGGAGCGCTCCATGATGCA 501
DB 2402 CGAGGAGCGCTCCATGATGCA 2422
RESULT 3
AB049841 2516 bp mRNA linear PRI 14-OCT-2000
LOCUS
AB049841 Macaca fascicularis brain cDNA, clone:QnpA-17096.
DEFINITION
AB049841
ACCESSION
AB049841.1 GI:10801619
VERSION

KEYWORDS fis (full insert sequence).
SOURCE Macaca fascicularis adult male brain parietal lobe cDNA to mRNA.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.

REFERENCE 1 (sites)
AUTHORS Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
Tero,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
TITLE Isolation of full-length cDNA clones from macaque brain cDNA
libraries

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2516)
AUTHORS Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-2000) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genbank/
Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)

COMMENT Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACTGCTG)
R. Site2: DraIII (CACTGCTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by Sugano et al. (University of Tokyo, Institute of
Medical Science). Custom primer used for sequencing
(5' end primer (CTTCTGCTTAAGCTGCG);
3' end primer (GGACCTGACCTGAGCACA)).

FEATURES
source
1..2516
Location/Qualifiers
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="Onpa-17096"
/sex="male"
/tissue_type="brain parietal lobe"
/clone_lib="macaque brain cDNA library Onpa"
/dev_stage="adult"
1242..>2516
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB16727.1"
/db_xref="GI:10801620"
/translation="MHHRIQQUGSLDNLSTFNHSHGSDADVLHOSLEANIATEVC
LTALDTLSLFLAKNQLADHGNPLMKVFVYVCFLOKHSEALNVPFALRPL
YKFPSTFEGRADMCALCEYLKCNKSLIRFASQLYFLMRNFVDYTKRSL
VHTLQVLIISOLIAADVIGIGTRFOOSLIINNCANSRLIKRTSFSSVDLTKR
ITVLMATAOKMEHNDPEMLVDLOYSLASVSTPELRKTRTLDMSARLHVNGDLTKR
AAMCYVHTALVAEYILTRKMGRCGCTAFVPIPNIDELSMEDVGMDDVHFNEDVL
MELLEOCADGLMKAREYELIADYIKLIPIPERPRFERLHLITLHRAYSKTEVEM
HSGRLDTTIRVAFRFGVSLSLVLOCPSFKE"

BASE COUNT 657 a 578 c 572 g 709 t
ORIGIN

Query Match 94.9%; Score 475.4; DB 9; Length 2516;
Best Local Similarity 96.8%; Pred. No. 6,6e-116;
Matches 485; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TTACACTGGAAAGAAGTCTTTGTCGGAACATTTGCAATCATATCTGTGACCA 60
DB 1700 TTACACTGGAAAGAAGTCTTTGTCGGAACATTTGCAATCATATCTGTGACCA 1759
QY 61 GCTGATGACAGAGTGTGTCGATTTGGGAAACAGATTCAGAGTCCGTGTCATCAT 120
DB 1760 GCTGATGACAGAGTGTGTCGATTTGGGAAACAGATTCAGAGTCCGTGTCATCAT 1819

QY 121 CAACAAGTGTGCCAGACAGTACCGGCTTATTAGACACACAGCTTCTCTGATGAA 180
DB 1820 CAACAAGTGTGCCAGACAGTACCGGCTTATTAGACACACAGCTTCTCTGATGAA 1879
QY 181 GCAGTTAACCAAAAGATACGACAGGTGCTAATGCGCACCGCCAGATGAAGACATGA 240
DB 1880 GCAGTTAACCAAAAGATACGACAGGTGCTAATGCGCACCGCCAGATGAAGACATGA 1939
QY 241 GAAGACCCAGACAGATGCTGTGACCTCCAGTACAGCCTGCGCAATTCATGTCACAC 300
DB 1940 GAAGACCCAGACAGATGCTGTGACCTCCAGTACAGCCTGCGCAATTCATGTCACAC 1999
QY 301 GCCGACCTCAGGAAGACGTGCTGCACAGCATGCGCCAGATTCATGCAAAAATGCGCA 360
DB 2000 GCCGACCTCAGGAAGACGTGCTGCACAGCATGCGCCAGATTCATGCAAAAATGCGCA 2059
QY 361 TCTCTCAGGACGACGAAATGCTATGCTACATACAGCCTGCTGAGAAATCTCTAC 420
DB 2060 TCTCTCAGGACGACGAAATGCTATGCTACATACAGCCTGCTGAGAAATCTCTAC 2119
QY 421 ACAGAAAGGCGTGTATGACAAAGATGACACCCCTTCAGGCTCATACCCCAACATGCA 480
DB 2120 ACAGAAAGGCGTGTATGACAAAGATGACACCCCTTCAGGCTCATACCCCAACATGCA 2179
QY 481 CGAGAGAGCCTCCATGATGCA 501
DB 2180 CGAGAGAGCCTCCATGATGCA 2200

RESULT 4
LOCUS AX255048 7506 bp DNA linear PART 10-OCT-2001
DEFINITION Sequence 7 from patent WO0170808.
ACCESSION AX255048
VERSION AX255048.1 GI:16074541
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 7506)
AUTHORS Rastelli,L.K. and Gerritsen,M.
TITLE Angiogenesis-associated proteins, and nucleic acids encoding the
same
JOURNAL Patent: WO 0170808-A 7 27-SEP-2001;
Curagen Corporation (US) : GENENTECH, INC. (US)
Location/Qualifiers
1..7506
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
1..7506
Location/Qualifiers

BASE COUNT 2147 a 1689 c 1764 g 1906 t
ORIGIN

Query Match 85.1%; Score 426.4; DB 6; Length 7506;
Best Local Similarity 99.8%; Pred. No. 8.1e-103;
Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTACACTGGAAAGAAGTCTTTGTCGGAACATTTGCAATCATATCTGTGACCA 60
DB 4572 TTACACTGGAAAGAAGTCTTTGTCGGAACATTTGCAATCATATCTGTGACCA 4631
QY 61 GCTGATGACAGAGTGTGTCGATTTGGGAAACAGATTCAGAGTCCGTGTCATCAT 120
DB 4632 GCTGATGACAGAGTGTGTCGATTTGGGAAACAGATTCAGAGTCCGTGTCATCAT 4691
QY 121 CAACAAGTGTGCCAGACAGTACCGGCTTATTAGACACACAGCTTCTCTGATGAA 180
DB 4692 CAACAAGTGTGCCAGACAGTACCGGCTTATTAGACACACAGCTTCTCTGATGAA 4751
QY 181 GGAATTAACCAAAAGATGACAGGCTGCTAATGCGCACCGCCAGATGAAGACATGA 240
DB 4752 GGAATTAACCAAAAGATGACAGGCTGCTAATGCGCACCGCCAGATGAAGACATGA 4811

QY 241 GAAGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCCTGGCCAAATCTATGCCAGCAC 300
|||||
Db 4812 GAAGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCCTGGCCAAATCTATGCCAGCAC 4871
QY 301 GCCGAGCTCAGGAAGACGTGGCTGCACAGCATGGCCAGAGATCATGTCCAAATAATGGCGCA 360
|||||
Db 4872 GCCGAGCTCAGGAAGACGTGGCTGCACAGCATGGCCAGAGATCATGTCCAAATAATGGCGCA 4931
QY 361 TCTCTCAGGACGACATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGATATCTCAC 420
|||||
Db 4932 TCTCTCAGGACGACATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGATATCTCAC 4991
QY 421 ACGGAAG 428
|||||
Db 4992 ACGGAAG 4999

RESULT 5
AF527605 7522 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens zizimln1 mRNA, complete cds.
DEFINITION AF527605
VERSION AF527605.1 GI:22038158
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7522)
AUTHORS Schwaartz,M.A., Iranli-Tehrani,M., Klosses,W.B., Del Pozo,M.A. and
TITLE zizimln1, a novel Cdc42 activator, reveals new guanine nucleotide
exchange-exchange factor domain for rho proteins
Nat. Cell Biol. (2002) In press
2 (bases 1 to 7522)
AUTHORS Meller,N. and Schwaartz,M.A.
TITLE Direct Submission
AUTHORS Submitted (05-JUL-2002) Cell Biology, The Scripps Research
JOURNAL Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA

FEATURES
source
1..7522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
56..6265
/note="guanine nucleotide exchange factor: Cdc42
activator"
/product="zizimln1"
/protein_id="AA080306.1"
/db_xref="GI:22038159"
/translation="MSQPLPLPASAETRTKTRALSKRPTAAELROSVSVVAGSVLLA
KPKLEPLDYENVIVOKTQILDLCREMLLPYDFOFALIRKRGRTICSTPAKAE
EEAQSLEFTECITKYNDSMDHLVNYKYEDYSGEFROLPNKVK.LDKLPVHYEDEV
KDEDAASLGSQKGITKHWLYKGNMNSAISVMRSKRRFFHLIOLGDSYMLNFK
DEKTSKEPKSGIFLDSMGVYONKVPAPFELKMOKRSYLLAAASEVMEWITIL
NKILOLNEAMOEKRNKDSHEDDESKESGSDGLSYPELAKASREKIKESST
RVKLEPLDPAOKLDFSSAEPEVKSFEKRGKRLKYCNLSINLCCCAENEEGPT
NVEPEFVTLDFDIKYNRKISADFHVDLNF SVROMLATTSPALMGSGOSPVYKGI
LHEAMQPKQKIFSVCPHEDIFLVARIEKVLQSGITTHCAEYKMSSDSKVAQVL
KNAKACORLGQYRMPFAMARLFLKASGNLKNARFSAIFYQDSNKSINDMLKVL
ADFRPEKMAKLPVILGNLDTIDNVSSDPNVNSYIPTKQFENCSKPTIFEEVE
FYPCPRKTOPRYTITNHLKYVPKYLKYDSOKSPAKRNALICIEFKDSDEDSOLK
CTYRGSGPVTRNSAFVYLHNHONPFYDEIKETLPQAHKHHLLPFHNSCNS
SKGSTRKRDVETGVISWPLLDGKRVVTSQITTPYSALMPSGIYGLQGLMGKRYG
PEIKWVDSKRLKISTHVLSTVYTOQHLHNFQYQCTQSGAQLAGELVATKLSL
HAMEGHVMIAPLITLNLQLEFVLTFRATOEVAVNTVYIHHVAQCHEELESHELSY
VYAYKAEVYASEKTYHEELTKSMITLIKPSADFLTSKPKYSEFEDYIKSMA
OHLTNSKRYKLNRQRPASVHAHVENYVMIMHITOKRPNDEAKNANHSILAVFI
KRCFTFMDRGVFKQINNYISCFAPGPKTLFEKKFELVVCNHEHYIPLN.PMPEG
KGRIRYDOLDDYSLDTEFCRNHFLVGLLREVTGLAOLQFREVRLAIVLKNLLIK
HSFDRVARSQARIATLYLFLGGLLENVORINAVDVSPFVNAQMTVXDSLAP
AVNPLVTPQKSTLDNSLHKDLGLAISGLIASPYTSTPINSVNAADSRGSLISTDSG

BASE COUNT 2146 a 1694 c 1774 g 1908 t
ORIGIN
Query Match 85.1%; Score 426.4; DB 9; Length 7522;
Best Local Similarity 99.8%; Pred. No. 8.1e-103;
Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAACCTGGAAGAAGTCTTGTGTCGGACACATTGTCAGTCAATCTCTGCAGCA 60
|||||
Db 4585 TTACACTGGAAGAAGTCTTGTGTCGGACACATTGTCAGTCAATCTCTGCAGCA 4644
QY 61 GCTGATAGCAGAGCTGTTGGCATTTGGGGAACGATTCACAGAGTCCCTGCATCAT 120
|||||
Db 4645 GCTGATAGCAGAGCTGTTGGCATTTGGGGAACGATTCACAGAGTCCCTGCATCAT 4704
QY 121 CACACACTGTGCGCAAGTACGCGGCTTATTAAGCACACACAGCTTCTCTGATGTGA 180
|||||
Db 4705 CACACACTGTGCGCAAGTACGCGGCTTATTAAGCACACACAGCTTCTCTGATGTGA 4764
QY 181 GGACTTAAACCAAAAGATTTAGCGACGCGGTAAAGGACACCGCCACAGATGAAGAGCATGA 240
|||||
Db 4765 GGACTTAAACCAAAAGATTTAGCGACGCGGTAAAGGACACCGCCACAGATGAAGAGCATGA 4824
QY 241 GAAGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCCTGGCCAAATCTATGCCAGCAC 300
|||||
Db 4825 GAAGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCCTGGCCAAATCTATGCCAGCAC 4884
QY 301 GCCGAGCTCAGGAAGACGTGGCTGCACAGCATGGCCAGAGATCATGTCCAAATAATGGCGCA 360
|||||
Db 4885 GCCGAGCTCAGGAAGACGTGGCTGCACAGCATGGCCAGAGATCATGTCCAAATAATGGCGCA 4944
QY 361 TCTCTCAGGACGACATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGATATCTCAC 420
|||||
Db 4945 TCTCTCAGGACGACATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGATATCTCAC 5004
QY 421 ACGGAAG 428
|||||
Db 5005 ACGGAAG 5012

RESULT 6
AB028981 7545 bp mRNA linear PRI 10-MAY-2002
LOCUS Homo sapiens mRNA for KIAA1058 protein, partial cds.
DEFINITION AB028981
VERSION AB028981.2 GI:20521745
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_ltb:plu:scrip1II SK plus
clone:hlh12146s1.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kikuno,R., Nagase,T., Ishikawa,K., Hirosewa,M., Miyajima,N.,
Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XIV. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 6 (3), 197-205 (1999)

MEDLINE 99397452
PUBMED 10470851
REFERENCE 2 (bases 1 to 7545)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@fokkazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
COMMENT On May 9, 2002 this sequence version replaced gi:5689452.
FEATURES
source Location/Qualifiers
1..7545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="h12146s1"
/tissue_type="brain"
/note="this sequence was obtained by subcloning of the DNA
fragments derived from two cDNA clones (1 - 1846 was
derived from ef00877 and 1847 - 7545 was derived from
h12146)."
1..7545
/gene="K1AA1058"
1..6288
/gene="K1AA1058"
/codon_start=1
/product="K1AA1058 protein"
/protein_id="BA83010.2"
/db_xref="GI:20521746"
/translation="LACICLARPPTGRSCSCLSLDLPFQPMQADKRTSSRSVKEL
VIESPLQKDAAGVEAESEPGVAPKLEIPDYENIVYOKTQIILNDCEKLEMLF
PYDDFOALIRROGRTICSTYPAKEDDEADSLSGCKGIRKGLYKGMNNSALSY
FROLNKKVYKIDKLPHVHYEVDDEVDDDEADSLSGCKGIRKGLYKGMNNSALSY
TMRSEKRRFPHLQIGDSYKNTNFKREKISKPEKSGISPLDSMGVYQNNKRRPFE
LMDKRSSTLLAASVEEMERWITLLKLIQLNEAMQKRNKRGDSDESKLEGS
GSGDSYSPLELAASAREAEIKLSESVKLIYLDPPAOKLIDESAPKESKLEFGK
RLIVCNDLSFNLQCVAAENEPPTNPEFFYLLSLFDIKYRKISAPFDHLMPS
VROMLATSIPALMNGSGSPGLKGIHEAMQYPRGIFSYCPDIFLVARLEK
LOGSITHCAPKESSDSKYAKVILNARAKOACORGOYMPAMAAATLFPDASGNL
DKNARESAIYRQDSNKLSDMDMLKLDAPKPRKMKLPVIGNIDITTDNYSDDPN
VNSSYITPKOETCSKPTPEVEEVEEPCIPKTHQPTTYITNHLVYKIKLYSOK
SPKARNIAICIEKEDDEEDSPLKCTIPGPGPYFTSAPFAVLAHNOEYDEI
KLEPQLHEKHHLILTFHYSCDSSKSGSTKRDVETQVGSMLPLKIDRVVTS
OHIPVANSIPSGYLGQYDGMGRHYGEIIMWVGKPLKISHLSTVYTOHILHN
FPOYQKTESGAQALGNELVYKLSLHMGHMLAFPLIILNQLRYVLTATOEHA
VAVTIVIIHVAQCHEEGLESLSRYKAYVKAEPYASQYKVEBELKSMITLIK
SADPLTSNKLKYSWEPFDVLIKMAQHLIENSKVLLRNORPAPASVHAHVVNML
MPTOKFRDNPASKNANSLAFITRCFTHPDRCGVPQIINNYISCTARPDKTLE
EKPEFLVYVNCHEHYIPLNPMPEFGKRIQRIODQLODLSLDECRNHLVGLILK
EYGILOERFVRLAISVLKNLIKHSFDRASRSHQRIATLPLFGLILNVQ
RLNVDVSPFPVNGMVTYKDESLAPVNPVLPQKSTLNDLSLHDLGALSGIASP
YITSTPNINSVFNADSRGSLISTDGSNLSPEKRSKNSLDKHOOSTGNSVNCCK
LDOSKISLMLCFYLIKSMEDDALFTYMKASTSEIMDEFTISEVCLHQFOYMKRY
IARTGMHARLQOGLSDNSLTFNHSYGHSDAVLHOSLLEANIATEVCLTDLTSL
FLAFKNOLADHGHNPLMKKVPDYLICFLOKHOSFALKNFTLARSILYEPSTFY
BERADMAICYEILKCCNSKTSITREASQLYLRNRPDGTGKSPVRHLYOII
SVSOLIAVYIGSTRQOSLITNNCANDRLIKHTSSDYKDLKTRIVLMAIA
OKHEHNDPEMLVDIOTSLAKSTASTELKRWDSMARIVANGULSADIAEYMHV
ALVAVETLRKRAVOWEPLPHSHSACLRSRGVFGGCTARVTPIMIDEASVME
DVGQDVHFNEDVLMLEOCADGLMAERYELADIYKLLPIYKRRKRLVRLAHLY
DTLIRAYSKVTEVMHSGRLIGTYRVAFGAQAQOTFSEIDVGFEDDEGKEYI
YKEPLTJLSEISORLLKLYSKDFGSENMKIDSGVKNKLDSDSYAIOYTHYEP
FDEKLEORKEFEKSHNIRFEMFETOTGRGSGVECKRRIILIAHCFYVYK
KRIPVYOHNDLNDIPALYALIDEMSKYVLEIOLCSAEDVMITOLKIGDSYVYNA
GLAYARAFLDDTNRKRPDKVYLLKEVFRQYERACGALANENLKEDELOEYOE
MKANREMAKLESEIMHEDLS"

Query Match 85.1%; Score 426.4; DB 9; Length 7545;
Best local Similarity 99.8%; Pred. No. 8.1e-103;
Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTACACTGGAAGAGCTTTTCCGGACATTTGCAAGTCANCTATCTGCACCA 60
Db 4608 TTACACTGGAAGAGAGCCCTTTGTCCGACACATTTGCAAGTCATATCTGCACCA 4667
QY 61 GCTGATAGCAGAGCTTTGTTGGCATTTGGGAAACAGATTCCAGAGTCCCTGTCATCAT 120
Db 4668 GCTGATAGCAGAGCTTTGTTGGCATTTGGGAAACAGATTCCAGAGTCCCTGTCATCAT 4727
QY 121 CAAACATGTGCCACACAGACACCGGCTTTATTAAGCACACACAGCTTCTCTGTATGGA 180
Db 4728 CAAACATGTGCCACACAGACACCGGCTTTATTAAGCACACACAGCTTCTCTGTATGGA 4787
QY 181 GGACTTTACCAAAAGATACGACAGGTGCTAATGAGCCAGCCAGATGAGAGAGCATGA 240
Db 4788 GGACTTTACCAAAAGATACGACAGGTGCTAATGAGCCAGCCAGATGAGAGAGCATGA 4847
QY 241 GAAGCACCACAGATGCTGTGTGACCTCCAGTACAGCCTGGCCAAATCTTATGCGACAC 300
Db 4848 GAAGCACCACAGATGCTGTGTGACCTCCAGTACAGCCTGGCCAAATCTTATGCGACAC 4907
QY 301 GCCCAGCTCAGAGACGACGTGCTGCACAGCATGGCCAGATCCATGTCAAAATGGCCGA 360
Db 4908 GCCCAGCTCAGAGACGACGTGCTGCACAGCATGGCCAGATCCATGTCAAAATGGCCGA 4967
QY 361 TCTCTCAGAGCAGCAATGTCTATGTCTCAGATACAGCCTAGTGGCAGATATCTCAC 420
Db 4968 TCTCTCAGAGCAGCAATGTCTATGTCTCAGATACAGCCTAGTGGCAGATATCTCAC 5027
QY 421 ACGGAAG 428
Db 5028 ACGGAAG 5035
RESULT 7
BC009134
LOCUS 2998 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, Similar to hypothetical protein FLJ20220, clone
MGC:11827 IMAGE:3596515, mRNA, complete cds.
ACCESSION BC009134
VERSION BC009134.1 GI:14318664
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2998)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gnathar, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAX Plate: 16 Row: 1 Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF

analysis, similarity but not identity to protein.

FEATURES

Source

Location/Qualifiers

1..2998

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:11827 IMAGE:3596515"

/tissue.type="Mammary tumor, C3(1)-tag model. Infiltrating ductal carcinoma, 5 month old virgin mouse."

/clone.lib="NCI_CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

381..1634

/codon_start=1

/product="Similar to hypothetical protein FLJ20220"

/protein_id="AAH09134.1"

/db_xref="GI:14318665"

/translation="MEROGCTAFRVITPNDIEASMMEDVGMQDVHFNEDVLMELLEQ

CADGLMKERYELADIKLITPIYERKDFERLAHLVYDLHRAYSKFEVHMSGRLL

LGTYPRVAFPGQGFEDDEKEXLYKEPKLTPLEISQRLKLSDFGSENVKMIQD

SGKVPKDLSDKFAYIOVTHVTFPEDEKELQERTEFERCHNIRREMFEMPTOTGR

OGVEEOCKRRTILTAIHCFYVKRIRPVMOHTDNLPLEVADMSKVALEHQLC

SSAEVDKIKDLQLOGSVYOVNAPLAYARAFDLDPNTRKYPDNKVKLKEFVFOFV

EACGQALVNERLIKEDOLYQEFEMKANYREMAKELSDIMREDICPLEKRTSVLPNSL

HIFNLSISTPISTYVQGLTSSSYV"

BASE COUNT 835 a 635 c 768 g 740 t

ORIGIN

Query Match 73.1%; Score 366.2; DB 10; Length 2998;
Best Local Similarity 88.2%; Pred. No. 8,5e-87;
Matches 398; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

51 CTGTGACCCAGCGATAGACAGCTGTTGGATTGGGGAACCAAGATCCAGCAGTCCC 110

1 CTGTGACCCAGCGATAGACAGCTGTTGGATTGGGGAACCAAGATCCAGCAGTCCC 60

111 TGTCCATCATCAACACTGTGCCACAGTACCGGCTATTATTAACACACAGCTTCTCT 170

61 TGTCTATCATCAACACTGTGCCACAGTACCGGCTATTATTAACACACAGCTTCTCT 120

171 CTGATGTGAAGACTTAACCAAAAGATACGACGCTGCTAATAGGCCACGCCCAATGA 230

121 CTGATGTGAAGACTTAACCAAAAGATACGACGCTGCTAATAGGCCACGCCCAATGA 180

231 AGGAGCATGAGACGACCGACAGATGCTGTTGACCTCAGTACAGCCTGGCCAAATCT 290

181 AGGAGCATGAGACGACCGACAGATGCTGTTGACCTCAGTACAGCCTGGCCAAATCT 240

291 ATGCCAGCAGCGCCGAGCTCAGGAAGAGCTGCTGCACAGCATGCGCAGATCTCA 350

241 ACGCCAGCAGCGCCGAGCTCAGGAAGAGCTGCTGCACAGCATGCGCAGATCTCA 300

351 AAATGSGCATCTCTCAGAGCAGCAATGTGCTATGTCACGTAACAGCCCTAGTGCAG 410

301 AAATGSGCATCTCTCAGAGCAGCAATGTGCTATGTCACGTAACAGCCCTAGTGCAG 360

411 AATATCTCAGACGGAAGCGTGTAAACAAGATGCAACCGCTTCAAGGTCATTAACC 470

361 AATATCTCAGACGGAAGCGTGTAAACAAGATGCAACCGCTTCAAGGTCATTAACC 420

471 CAACAATCGACGAGAGGCGCTCCATGATGA 501

421 CAACAATCGATGAAGAGGCTTCATGATGA 451

RESULT 8

LOCUS R. norvegicus tlg mRNA. 3227 bp mRNA Linear ROD 24-FEB-1999

DEFINITION X68101

ACCESSION X68101.1 GI:550419

VERSION tlg gene.

KEYWORDS Rattus norvegicus.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

AUTHORS

PIANESSE, L.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

AUTHORS

PIANESSE, L., PORCELLINI, A., AVVEDIMENTO, V.E., D'ESPOSTI, F.,
FELICIELLO, A., MONTICELLI, A., MUSTI, A.M., TORTORA, G., VARRONE, S.
and Cocozza, S.

TITLE

JOURNAL

REFERENCE

QY 362 CTCCTCAGAGCAGCATGTGCTATGTCCACGATGAGCCCTAGTGAGCAATATCTCACA 421
|||||
Db 635 CTCCTCAGAGCAGCATGTGCTATGTCCACGATGAGCCCTAGTGAGCAATATCTCACA 694
QY 422 CGGAAG 428
|||||
Db 695 CGGAAG 701
RESULT 9
AX058192 2610 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 62 from Patent WO0077040.
ACCESSION AX058192
VERSION AX058192.1 GI:12310693
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2610)
AUTHORS Yue,H., Tang,Y.T., Hillman,J.L., Lai,P., Bandman,O., Baughn,M.R.,
Azizmai,Y., Yang,J., Reddy,R. and Lu,D.A.
TITLE Human Intracellular signaling molecules
JOURNAL Patent: WO 0077040-A 62 21-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..2610
BASE COUNT 878 a 467 c 529 g 736 t
ORIGIN
Query Match 54.8%; Score 274.4; DB 6; Length 2610;
Best Local Similarity 71.8%; Pred. No. 2.5e-62;
Matches 359; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
QY 2 TACACTGGAAGAGTCTTGTTCGCGACACATTTGCAAGTCATCATATCTGTCAGCAG 61
|||||
Db 591 TATACCAAGAGAAACCTTTTGAGGACACATCTACAGATAATATTTGCTGAAGCAA 650
QY 62 CTGATAGCAGAGCTTTGGCATTTGGGAAACGAGATTTCCAGCAGTCCCTGTCATCATC 121
|||||
Db 651 CTGATAGCTGATGTAGCAGCTAAGCGAGGATCAGAGATTTTCAGAGCTTTATTCATATATC 710
QY 122 AACAACTGTGCAACAGTGAACCGGCTTATTAAAGCACACAGCTTCTGCTGATGGAAG 181
|||||
Db 711 AATATATTTGCAATAGTACAGACACATATGAAGCACTGCTTTCCCGCAGAGTCAAA 770
QY 182 GACTTAAACCAAGATACGACGCTGCTAATGCGCACCGCCAGATGAAGAGCATGAG 241
|||||
Db 771 GACTTGAAGCAAGAAATCCGACATGTTCTATGCGCACCTCCCAATGAAGAGCATGAG 830
QY 242 AACACCCAGATGCTGTGAGCTCCAGTACGCTTGGCCAAATCTTATGTCAGCAGC 301
|||||
Db 831 AAACACCTGAAATGCTAATTTGATCTCCAGATAGCTTACCAAGTCTTATGCAAGCACC 890
QY 302 CCGAGCTCAGAGAGAGCTGCTGACAGACATGCGCAGATGCCATGCAAAAATGGCGAT 361
|||||
Db 891 CCAAGAGCTCAGAGAAACCTGCTTGAATGATGCGCAAGTTTCATGTAATAAATGGAGAT 950
QY 362 CTCCTCAGAGCAGCATGTGCTATGTCCACGATGAGCCCTAGTGAGCAATATCTCACA 421
|||||
Db 951 TTTTCAGAGGCTGAGATGTATTATGTCATAGCAGACCTAGTGTGAGAGTTTCTTCAT 1010
QY 422 CGGAAGGCTGTTTAAACAAGATGACCGCTTCAGGCTCATTTCCCAACCAATCGAC 481
|||||
Db 1011 CGAAGAAATTTATTTCTTACAGATGTTTCAGGCTTCAAGAAATTTACTCCCAATATAGAT 1070
QY 482 GAGGAGCTTCATGATGA 501
|||||
Db 1071 GAAGAGGAGCAATGAAGA 1090

RESULT 10
AK056684
LOCUS
DEFINITION Homo sapiens cDNA FLJ32122 fis, clone PEBLM1000144, moderately similar to Tr9.
ACCESSION AK056684
VERSION AK056684.1 GI:16552156
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens peripheral blood mononuclear cells (HPBMC5939) cDNA to mRNA, clone_11b:PEBLM1 clone:PEBLM1000144.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Hotuta,T., Hiraoaka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chida,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
2 (bases 1 to 3152)
REFERENCE Isogai,T., Otsuki,T. and Sugiyama,T.
AUTHORS Direct Submission
TITLE Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
JOURNAL (E-mail:genomices@nrl.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-8 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; RAB and HRI.
FEATURES
source location/Qualifiers
1..3152
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PEBLM1000144"
/cell_type="peripheral blood mononuclear cells
(HPBMC5939)"
/note="PEBLM1"
/note="cloning vector: pME18SFL3-Primary culture, peripheral blood mononuclear cells"
154..2799
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71253.1"
/db_xref="GI:16552157"
/translation="MPNSASDPDEPCGCTSPANGSISTDKTAYVGFONGHGTAKRD
SRGSLPEGATGCFPDGNTGENTROSTRSSVSOTNRIDQETISLMLCTLYIVKMS
EDTLITYNKVSPOELNITLILEVCLHEFYMKRINARVHADMLSKHGIDKRSOT
MPLENRSGVMQARLOHLSLSESETLNHSSTTTEADIFQALLEGTAIEVSLTVD
TISEPTQCFKQLOLNNDGNHDMKRVDFIHIAFLKNGSEVSLKHVPASRAFLSKRP
SAFPGRVNMCAPCYEVLKCTSKISTREASALILYLMRNNFEYTRKTRPRL
OITAVSOLINDVIALSGSROESLFTIINNANSDRPMKATAPAEVKDITKRTRYL
MATAMCKHEMDPEMLDLOYSLAKSVASIPDELKTUWDSAKIIVANGSPSEAMKY
VYVAALVAEPFLHRKMLPENGSAFKITLPINDEGAKAGMADVHYSEVLELLE
OCVGLMVAERYELIISLSKILVPIYKRREBEKLDIVYTLGALVYKILEVNHKTR
LIGTFEFAVFGOSFEEDGKEVETPEKLETKETFEENHNSIRFVEAPATLTKK
LSDKVNAKELDPKTAHIOVTVKPYFDEKLETKETFEENHNSIRFVEAPATLTKK
KGGCIEQCKRRTILUTSNSPPYKRIPIINGEQINLRIDVATDEIKKTKYELDKL
CSSNDVMIOLOLILGDCVSYOVAGGLATAPARFLNDSQASKIPPKVSELEDMFRKF
IQACSIATELNERLIKEDOVETHEGLSNRDMVKELSDIITHBOILIQEDTMSHPMSN
TLHFVCAISGTSIDRGYGSPPYAV"

BASE COUNT

1058 a 576 c 638 g 880 t


```
RESULT 12
AX173120
LOCUS AX173120 3472 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 99 from Patent WO0142294.
ACCESSION AX173120
VERSION AX173120.1 GI:14598013
KEYWORDS
SOURCE
.
ORGANISM
synthetic construct.
artificial construct.
REFERENCE
1 (bases 1 to 3472)
AUTHORS
Lu, P., Garman, J.D. and Candia, A.F.
TITLE
Clasp-4 transmembrane protein
JOURNAL
Patent: WO 0142294-A 99 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
Location/Qualifiers
source
1..3472
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"
BASE COUNT 1153 a 640 c 696 g 983 t
ORIGIN
Query Match 53.8%; Score 269.6; DB 6; Length 3472;
Best Local Similarity 71.2%; Pred. No. 4.8e-61;
Matches 356; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 2 TACACTGGAAAGAGTCCCTTTGTCGGACACATTTGCAAGTCATCATATCTGCAGCCAG 61
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1461 TATACCAAAAGGAAACCTTTTGAGGACACATCTAGATAATTAATGCTGAAGCAA 1520
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 62 CTGATGACAGACGTTGTTGGCATGGGGAACCAAGATTCACAGACGCCCTGCATCATC 121
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1521 CTGATGACGATGTGACACCTAAGGAGATCAAGATTTACAGAGCTCTTATTCATATATC 1580
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 122 AACACGTGTGCCAACAGTACCGGCTTATTAAACACACCAAGCTTCTCTCTGATGTGAAG 181
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1581 AATAATTTTGGCAATAGTGCAGACACTATGTTGGCAAGACGCTTTCGCCGAGAAGTCAAA 1640
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 182 GACTTAACCAAAAGAGATGACGAGCTGTAAAGGCCACCCGCCACAGATGAAGAGCATGAG 241
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1641 GACTTGAACCAAGAAATCCGACTGTTCTTATGCGCACCTCCCAAAAGGAAGAGCATGAG 1700
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 242 AACGACCCAGAGATGCTGTGTGACCTCAGTACAGCCTGSCCAATGCTAATGCGAGCAG 301
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1701 AAAGACCTGGAATGCTAATGATCTCCAGTATAGCTTACCAAGTCTCTATGCAAGCACC 1760
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 302 CCCGAGCTCAGGAAGAGAGCTGTGACAGCAGTGCAGAGATTCATAAAAATGGCGAT 361
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1761 CCAGAGCTCAGGAAGAACTGTGATGATGCGCAAGATCATGTAATAAATGAGAT 1820
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 362 CTCACAGAGGAGAGATGCTATGTCACAGTAAGAGCCCTAGTGCAGATATCTCACA 421
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1821 TTTTCAAGAGCTGCGATGATGTTATGTCATGATGACAGCTCTAGTTCAGAGTTTCTTCAN 1880
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 422 CGGAAGAGCGTGTATTAGACAGAGTGCACCGCTTCAGGGTCTATTACCCCAACATGCAG 481
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1881 CGAATAAATAATTTTCTTAAGCGATGTTCAAGCGTTCAAGAAATTTACTCCCAATATAGAT 1940
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 482 GAGGAGGCTTCATGATGGA 501
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1941 GAAGAAAGAGCAATGAAGA 1960
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
RESULT 13
AX173028
LOCUS AX173028 4391 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 7 from Patent WO0142294.
ACCESSION AX173028
VERSION AX173028.1 GI:14597978
KEYWORDS
```

```
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 4391)
AUTHORS
Lu, P., Garman, J.D. and Candia, A.F.
TITLE
Clasp-4 transmembrane protein
JOURNAL
Patent: WO 0142294-A 7 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source
1..4391
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="human CLASP-4 cDNA"
414..4058
/note="human CLASP-4"
/codon_start=1
/protein_id="CAC43730.1"
/translation="MEIQVLRFLSVILMOLFVYLPWIMHEDDVISCWPVLIPIVSK
CHEEGIDYLSLSEFIKYSFRPKPSAPQAPLHETLATMMLALLKQSDPLAINKLKY
SWPEFELIAKSMATYLLLEENKIKLPHQGRPKAYHMLHLSFLAATTVESQVATPRE
SRNVVYSLASFLLKCCLTLMDRGFVENLINDYISQSPKDPVLALEFEELOTICNHE
HYIPLNLFMAFAKPKLQROVDNLEYSLSDEYKHFHVLGLLEIETSLADQNEYIRY
TAISYIKNLILKHAFFDTRYOKHNOQAKIAQVLPVGLLENIGRLAGDTIYSCAM
PNSASRDEFPFGFTSPANRGLSTDQKTAVSQPNHGKIKREDRSGLIEGATGFPD
QGNNGENRQSTSSVSQYRILQVETIRSLMCLYLYTKNISDITLLTYWNVSPDE
LINITILLEVCLFHRVYMKRNRIARVHDAMVLSKHFGLDRSQSOTPALRNNSQWQAL
QHLSESLFTLNHSTTTEADIFPHQALBENFTPEVSLVLDITISFEQCFPHFEN
NDGHPMLKRVFDIHLAFKNGQSEVSLKHFVAFSLRPFISKPSAFKGVNMAAC
YEVLKCTSKLSSTRNEASALLYLMMRNPEYTRKTELRLHQLITIAVSQILADVL
SGSRSFOCSLTILINFNASDRPMLARAPFAVVKDITRKIRITVIMATQOMEHEHEDPM
LIDLOYSLAKSYASTPELRKTYWDSMAKIRHKNDSFSAACVYHAVALVAFELHRK
LFPNGSAFKKITFRNIDEGAMKDDAGMDVHSSEVLELEECVNGLMAEYETI
SEISKLIGPIYENRREFENLQVRYRLHGAATKILEVMHTRKRLGLTFEYVAYEGOSF
FEEDBGEKEYIKYKEBLTGLSEISLRLVLYKGEKGTENVXIIQSDVKNKKEIDPKYA
HIOVYVYKPYDPDEKELERKTEFERNNHISREVEAPYTLGSKOGCIEBOCKRRTLU
TTSNSFVYKRRIPINCEQOINLRFIDGATDEIKDFAELQKLCSSSTDVMIQDLRL
QGWVSVQVNAGLAVYARAFILNDSQASVPPKVSSELMKDRKTPQACSTALEINERLI
KEDQVEYHEGILKSNFRDMVKRLSIIHBOILQEDTHMSPWMSNTLHVFCATISGTSDBR
GYGSRVYAEV"
BASE COUNT 1359 a 851 c 908 g 1273 t
ORIGIN
Query Match 53.8%; Score 269.6; DB 6; Length 4391;
Best Local Similarity 71.2%; Pred. No. 4.8e-61;
Matches 356; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 2 TACACTGGAAAGAGTCCCTTTGTCGGACACATTTGCAAGTCATCATATCTGCAGCCAG 61
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2379 TATACCAAAAGGAAACCTTTTGAGGACACATCTACAGATAATTAATGCTGAAGCAA 2438
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 62 CTGATGACAGACGTTGTTGGCAATTGGGAAACCAAGATTCACAGAGTCCCTGTCATCATC 121
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2439 CTGATGACGATGTGACACTAAGCCGAGATCAGAATTTCAAGAGTCTTTATTCATATATC 2498
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 122 AACCACTGGCCACAGACGCGGCTTATTAAGCAGACACAGTCTCTCTGATGTGAG 181
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2499 AATAATTTTGGCAAAATAGTGACAGACCTTATGTTGCCAAGAGCTCCCGCAGAAAGTCAA 2558
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 182 GACTTAACCAAAAGAGATACGACGAGTGTAAATGGCCACGCCAGATGAAGAGCATGAG 241
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2559 GACTTGACCAAGAGAAATCCGACGCTGTTCTTATGGCCACTGGCCAAATGAAGAGCATGAG 2618
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 242 AACGACCCAGAGATGCTGTGACCTCCAGTACAGCTGTGCGCAATTCCTATGCAAGCAG 301
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2619 AAAGACCTGGAATGCTAATGATCTCCAGTATAGCTTAAGCCAAGTCTTATGCAAGCACC 2678
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 302 CCCGAGCTCAGGAAGAGCGGCTGCAGACAGTGCAGAGATCCATGCAAAAATGGGAT 361
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2679 CCAGAGCTCAGGAAGAACTGGCTTATGACATGGCCAGATTCATGTAAATAAATGAGAT 2738
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 362 CTCACAGAGGAGCAATGCTATGTCCACGTAACAGCCCTTAGTGGCAGAAATATCTCACA 421
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
```

Db	2739	TTTTAGAGAGGCTTCCATGCTTATATGCTATGACAGGCTTACGAGCTTCTTCAT	2798
QY	422	CGGAAAGGCGTGTTTAGACAAAGATGCAACCGCTTCAGGGCTCATTAACCCAAACATCGAC	481
Db	2799	CGAAAAAATTTTCTCTTAACGAGATGTCAGCGTTCAAGAAAAATTAATCTCCAAATATGAT	2858
QY	482	GAGGAGGCTCCATGATGCA 501	
Db	2859	GAGGAGGAGCAATGAAAGA 2878	
<p>RESULT 14</p> <p>AX173118 4393 bp DNA linear PAT 03-JUL-2001</p> <p>LOCUS AX173118</p> <p>DEFINITION Sequence 97 from Patent WO0142294.</p> <p>ACCESSION AX173118</p> <p>VERSION AX173118.1 GI:14598012</p> <p>KEYWORDS</p> <p>SOURCE</p> <p>ORGANISM</p> <p>REFERENCE</p> <p>AUTHORS</p> <p>TITLE</p> <p>JOURNAL</p> <p>FEATURES</p> <p>source</p> <p>1..4393</p> <p>/organism="synthetic construct"</p> <p>/db_xref="taxon:32630"</p> <p>/note="polynucleotide fragment"</p> <p>RASE COUNT 1359 a 852 c 910 g 1272 t</p> <p>ORIGIN</p> <p>Query Match 53.8%; Score 269.6; DB 6; Length 4393;</p> <p>Best Local Similarity 71.2%; Pred. No. 4,8e-61;</p> <p>Matches 356; Conservative 0; Mismatches 144; Indels 0; Gaps 0;</p>			
QY	2	TACACTGCAAAAGAGCTTTGTCCGACACATTTCCAGTCATCATATCTGTCAGCCAG	61
Db	2379	TATACCAAAAGAAAACCTTTTGGAGACACATCTACATATATATTTCTGTAAAGCAA	2438
QY	62	CTGATAGCAGACGTTGTTGGCATTTGGGGAACACAGATTCCAGCAGTCCCTGTCCATATC	121
Db	2439	CTGATAGCTGATGTAGCACTAAGCGAGGATCAAGATTTCAGGAGTCTTATTCATTATC	2498
QY	122	AACCACTGTGCCAAGTACAGTACCGGCTTATTAAGCACACCAAGCTTCTCTGTATGGAAG	181
Db	2499	AATTAATTTTGGCAAAATAGTACAGACCTATGTTGGCAAGAGCCTTTCCCCAGAAAGTCAAA	2558
QY	182	GACTTAACCAAAAAGGATAGCAGCGGTGCTAATGTGGCCACCGCCAGATGGGAAGAGCATGAG	241
Db	2559	GACTTGTACCAAGAGATTCGCACTGTTCTTATGTGGCCACTGCCCCAAATGAAAGAGCATGAG	2618
QY	242	AACGACCCAGAGATGCTGTGAGCACTCCACAGTACAGGCTGCGCCAAATCTATGCGACAGC	301
Db	2619	AAAGACCTCTGAATGTATATGATCTCCAGTATAGCTTGAAGCAAGTCTATGCAAGCACAC	2678
QY	302	CCGAGACTAGGAAGACGTGCTCCAGACAGCATGCGCAGATCCATGTCAAAAATGGCGAT	361
Db	2679	CCAGAGCTCAGGAAAACCTGGCTGTATGACATAGGCCCAAGATTCATGTAAAAATGGAGAT	2738
QY	362	CTCTCGAGAGCAGCATGTGCTATGTCACAGTAACAGAGCCCTGAGGCAAAATATCTCAAA	421
Db	2739	TTTTTGAAGGCTGCGATGTGTATGTCTCCATGTAGCAGCTCTAGTTCACAAGTTTCTCAT	2798
QY	422	CGGAAAGGCGTGTTTAGAACAGGATCACCGGCTTCAGGCTATTAACCCCAACATCGAC	481
Db	2799	CGAAAAAATTTTCTCTTAACGAGATGTCAGCGTTCAAGAAAAATTAATCTCCAAATATGAT	2858
QY	482	GAGGAGGCTCCATGATGCA 501	

Db	2859	GAAGAGAGCAATGAAGA	2878
RESULT 15			
AK054649		2299 bp	linear
LOCUS			
AK054649			
DEFINITION			
AK054649			
VERSION			
AK054649.1			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 359;			
Conservative			
2			
287			
62			
347			
122			
407			
182			
467			
52.6%			
71.7%			
0;			
Mismatches 141;			
Indels 1;			
Gaps 1;			
Location/Qualifiers			
1..2299			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="BNHG41000003"			
/cell_line="H4"			
/cell_type="neurogloma"			
/clone_lib="BNHG41"			
/note="cloning vector: PME18SFL3"			
780 a			
392 c			
471 g			
656 t			

QY 242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCTATGCCAGCAG 301
|| |||| || |||| || || |||| || || |||| || || |||| || || ||||
Db 527 AAGACCCCTGAATGTATGTATGTCTCCAGATATAGCTTAGCCAGAGTCTATGCAGAGACC 586
QY 302 CCCGAGCTCAGAGAGACGTGGCTCGACAGCATGGCCAGAGATCCATGTCAAAAAATGGCGAT 361
|| |||| |||| || |||| || || |||| || || |||| || || |||| || || ||||
Db 587 CCAGAGCTCAGAGAAAAACCTGGCTGTAGCATGGCCAGAGATTCAATGTAAAAAATGGAGAT 646
QY 362 CTCTCAGAGGACAGCATGTGTATGTCCAGTACAGGCCCTAGTGGCAGATATTC-TCAAC 420
| |||| || || |||| |||| || || || |||| || || |||| || || ||||
Db 647 TTTTCAGAGGCTGCCGATGTGTATGTCCATGTAGCAGGCTCTAGTGCAGAGTTCTTCAT 706
QY 421 ACGGAAAGGCGTGTAGACAGAGATGCACCGCCTTCAGGCTCATTAACCCCAACATCGA 480
|| || || || || || || || || || || || || || || || || || || || || ||
Db 707 CGAAAAAATTTATTTCTTAACGAGATGTTCAAGCGTTCAAGAAAAATTACTCCCAATATAGA 766
QY 481 CGAGAGGCCCTCCATGATGA 501
|| || || || || || || || || || || || || || || || || || || || || ||
Db 767 TGAAGAGAGGACATGAAGA 787

Search completed: February 7, 2003, 08:03:39
Job time : 1021.66 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:36:04 ; Search time 128.085 Seconds
(without alignments)
8808.598 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000
Perfect score: 501
Sequence: 1 ttacactggaagaagtcct.....gagagagcctccatgatgga 501

Scoring table: IDENTITY_NMC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	3614	24	ABK84967 DNA encoding cadhe
2	501	100.0	3642	24	ABK84971 DNA encoding cadhe
3	501	100.0	3705	24	ABK84968 DNA encoding cadhe
4	501	100.0	4806	24	ABK84964 DNA encoding cadhe
5	501	100.0	4807	21	AACT87972 Human CLASP-2 nucl
6	501	100.0	4807	21	AACT87973 Human CLASP-2A nuc
7	501	100.0	4807	24	ABK84966 DNA encoding cadhe
8	501	100.0	4807	24	ABK84973 DNA encoding cadhe
9	501	100.0	4898	21	AACT87974 Preliminary CLASP-

10	501	100.0	4898	21	AACT87975 Preliminary CLASP-
11	501	100.0	4898	21	AACT87976 Preliminary CLASP-
12	501	100.0	4898	21	AACT87977 Preliminary CLASP-
13	501	100.0	4898	21	AACT87978 Preliminary CLASP-
14	501	100.0	4898	21	AACT87979 Preliminary CLASP-
15	501	100.0	4898	21	AACT87980 Preliminary CLASP-
16	501	100.0	4898	21	AACT87981 Preliminary CLASP-
17	501	100.0	4898	21	AACT87982 Preliminary CLASP-
18	501	100.0	4898	24	ABK84992 DNA encoding cadhe
19	501	100.0	4898	24	ABK84993 DNA encoding cadhe
20	501	100.0	4898	24	ABK84994 DNA encoding cadhe
21	501	100.0	4898	24	ABK84995 DNA encoding cadhe
22	501	100.0	4898	24	ABK84996 DNA encoding cadhe
23	501	100.0	4898	24	ABK84997 DNA encoding cadhe
24	501	100.0	4898	24	ABK84998 DNA encoding cadhe
25	501	100.0	6791	24	ABK85003 DNA encoding cadhe
26	499.4	99.7	3678	23	ABV21145 DNA encoding cadhe
27	499.4	99.7	3899	22	AAF32675 Human prostate exp
28	426.4	85.1	5862	24	ABK84970 Human CDNA encodin
29	426.4	85.1	6816	21	AACT74524 Human CDNA encodin
30	426.4	85.1	7506	22	AACT91118 Human OREFX ORF79 p
31	422	84.2	5048	24	AACT32647 Angiogenesis assoc
32	274.4	54.8	2610	22	AACT32647 Human CDNA encodin
33	274.4	54.8	6454	22	AACT32647 Human CDNA encodin
34	269.6	53.8	3472	22	AACT32647 Human CDNA encodin
35	269.6	53.8	4393	22	AACT32647 Human CDNA encodin
36	221.2	44.2	5214	21	AACT87968 Human CDNA encodin
37	221.2	44.2	5214	21	AACT87968 Human CDNA encodin
38	216.4	43.2	5688	21	AACT87969 Mouse CLASP-1 nucl
39	216.4	43.2	5688	21	AACT87969 Human CDNA encodin
40	216.4	43.2	7277	24	AACT87969 Human CDNA encodin
41	171.4	34.2	5955	23	ABCT11149 Human CDNA encodin
42	171.4	34.2	8942	23	ABCT11148 Human CDNA encodin
43	168.2	33.6	1476	24	AACT2670 Human CDNA encodin
44	136	27.1	2120	22	AACT2670 Human CDNA encodin
45	130.6	26.1	15057	22	AACT2670 Human CDNA encodin

ALIGNMENTS

RESULT 1	ABK84967	standard; cDNA; 3614 BP.
ID	ABK84967	
AC	ABK84967;	
XX		
DT	13-AUG-2002 (first entry)	
XX		
DE	DNA encoding cadherin-like asymmetry protein (CLASP) isoform #2.	
XX		
KW	Human; autoimmune disease; hematopoietic disorder; DiGeorge syndrome;	
KW	blood protein disorder; agammaglobulinemia; dysgammaglobulinemia;	
KW	ataxia telangiectasia; common variable immunodeficiency; lymphoma;	
KW	thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;	
KW	haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;	
KW	endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;	
KW	autoimmune pulmonary inflammation; organ rejection; inflammation;	
KW	CLASP; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	W020023117-A2.	
XX		
PD	18-APR-2002.	
XX		
PF	15-OCT-2001; 2001WO-US32202.	
XX		
PR	13-OCT-2000; 2000US-0687837.	
XX		
PA	(ARBO-) ARBOR VITA CORP.	
PA	(GARM/) GARMAN J D.	
PA	(CAND/) CANDIA A F.	

```
XX
PI Lu PS:
XX
DR WPI: 2002-416861/44.
DR P-PSDB: ABG61673.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis
XX
PS Disclosure: Figure 3A: 245bp: English.
XX
XX The invention relates to an isolated polypeptide (I) comprising an amino
XX acid sequence that has 90 % sequence identity to one of the human
XX cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
XX sequences (PS). (I) is useful for identifying a compound or agent that
XX binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
XX detecting a CLASP-2 polypeptide in a sample. (III) is useful for
XX inhibiting a immune response in a subject. A pharmaceutical composition
XX comprising a nucleic acid encoding (I), or (II) is useful for preventing
XX or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
XX the autoimmune disease is caused or exacerbated by increased activity
XX of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
XX detection or inhibition of CLASP-2 expression (e.g., antisense or
XX ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
XX polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
XX antibodies or are used as therapeutic polypeptides. The CLASP-2
XX polynucleotide or fragments can be used in diagnostics (e.g., as probes
XX for CLASP-2 expression), as a lymphocyte marker and for therapeutic
XX purposes. CLASP-2 polynucleotides can construct transgenic and knockout
XX animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
XX polynucleotides can screen for CLASP-2 agonists and antagonists.
XX CLASP-2 polypeptides or polynucleotides can treat deficiencies or
XX disorders of the immune system, by activating or inhibiting the
XX activation, differentiation of immune cells and can treat or detect
XX deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
XX or polynucleotides can increase differentiation and proliferation of
XX haematopoietic cells, including the pluripotent stem cells to treat those
XX disorders associated with a decrease in certain (or many) types of
XX haematopoietic cells (e.g., immunologic deficiency syndromes including
XX blood protein disorders (e.g., agammaglobulinemia,
XX dysgammaglobulinemia, ataxia telangiectasia, common variable
XX immunodeficiency, DiGeorge syndrome, lymphopenia, thrombocytopenia, or
XX haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
XX detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
XX Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
XX endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
XX inflammation. CLASP-2 can be used to treat anaphylaxis or
XX hypersensitivity to an antigenic molecules, organ rejection or graft-
XX versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
XX represent cadherin-like asymmetry protein (CLASP) coding sequences and
XX PCR primers of the invention.
XX
SQ Sequence 3614 BP: 1009 A; 834 C; 874 G; 897 T; 0 other;
XX
Query Match 100.0%; Score 501; DB 24; Length 3614;
Best Local Similarity 100.0%; Pred. NO. 9.9e-142;
Matches 501: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TTACACTGGAAGAGAGTCTTTGTCGGACATTTGCAGTCATCATATCTGTCAGCA 60
DB 1307 TTACACTGGAAGAGAGTCTTTGTCGGACATTTGCAGTCATCATATCTGTCAGCA 1366
QY 61 GCTGATAGCAGAGCGTTGTTGGCATTGGGAAACAGATTCAGCAGTCCTGTCATCAT 120
DB 1367 GCTGATAGCAGAGCGTTGTTGGCATTGGGAAACAGATTCAGCAGTCCTGTCATCAT 1426
QY 121 CAACACGCTGCGACAGTACCGGCTTATTAAGACACCGCGCTTCCTGATGTTAA 180
DB 1427 CAACACGCTGCGACAGTACCGGCTTATTAAGACACCGCTTCCTGATGTTAA 1486
QY 181 GGACTTAACCAAGATACGACGCTGCTAATGCGCACCGCCAGATGAAGAGCATGA 240
```

```
DB 1487 GGACTTAACCAAGATACGACGCTGCTAATGCGCACCGCCAGATGAAGAGCATGA 1546
QY 241 GAACGACCCAGAGATCGTGGTGACCTCCATGACAGCTGCGCAATCTTATGCCAGAC 300
DB 1547 GAACGACCCAGAGATCGTGGTGACCTCCATGACAGCTGCGCAATCTTATGCCAGAC 1606
QY 301 GCCCGAGCTCAGAGAGACGTGGCTGCACAGCATGGCCGAGATCCATGTCAAAATGGGCA 360
DB 1607 GCCCGAGCTCAGAGAGACGTGGCTGCACAGCATGGCCGAGATCCATGTCAAAATGGGCA 1666
QY 361 TCTCTAGAGAGCAGCATGTGCTATGTCACGCTAACAGCCCTAGTGGAGAAATATCTCAG 420
DB 1667 TCTCTAGAGAGCAGCATGTGCTATGTCACGCTAACAGCCCTAGTGGAGAAATATCTCAG 1726
QY 421 ACGGAAAGCGCTGTTTAGACAAAGATGACCGCCCTTCAGGTCATATTACCCAAACATGGA 480
DB 1727 ACGGAAAGCGCTGTTTAGACAAAGATGACCGCCCTTCAGGTCATATTACCCAAACATGGA 1786
QY 481 CGAGGAGCGCTCCATGATGGA 501
DB 1787 CGAGGAGCGCTCCATGATGGA 1807

RESULT 2
ABK84971
ID ABK84971 standard; cDNA; 3642 BP.
XX
AC ABK84971:
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #6.
XX
XX Human: autoimmune disease; haematopoietic disorder; DiGeorge syndrome;
XX blood protein disorder; agammaglobulinemia; dysgammaglobulinemia;
XX ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
XX thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
XX haemolytic anemia; multiple sclerosis; rheumatoid arthritis; lupus;
XX endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
XX autoimmune pulmonary inflammation; organ rejection; inflammation;
XX CLASP; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS:
XX
DR WPI: 2002-416861/44.
DR P-PSDB: ABG61677.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis
XX
PS Disclosure: Figure 3A: 245bp: English.
XX
XX The invention relates to an isolated polypeptide (I) comprising an amino
XX acid sequence that has 90 % sequence identity to one of the human
XX cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
XX sequences (PS). (I) is useful for identifying a compound or agent that
```

CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting an immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g., an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, DiGeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecule, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK4922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.

XX Sequence 3642 BP; 1005 A; 899 C; 877 G; 861 T; 0 other;

SQ Query Match 100.0%; Score 501; DB 24; Length 3642;

Best Local Similarity 100.0%; Pred. No. 9, 9e-142;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACACTGGAAAGAGTCCCTTTCGACACATTTCAGATCATATCTGACCCA 60
DB 2028 TTACACTGGAAAGAGTCCCTTTCGACACATTTCAGATCATATCTGACCCA 2087
QY 61 GCGATAGCAGAGTGTGGCATTGGGAAACAGATTCCAGAGTCCCGTCCATCAT 120
DB 2088 GCGATAGCAGAGTGTGGCATTGGGAAACAGATTCCAGAGTCCCGTCCATCAT 2147
QY 121 CAACAAGTGTGCCAAGTACCGGCTTATTAGACACACAGCTTCTCCTGTGATGAA 180
DB 2148 CAACAAGTGTGCCAAGTACCGGCTTATTAGACACACAGCTTCTCCTGTGATGAA 2207
QY 181 GGACTTAACCAAAAGATAGCAGCGTCTAATATGCCACCGCCAGATGAGAGCATGA 240
DB 2208 GGACTTAACCAAAAGATAGCAGCGTCTAATATGCCACCGCCAGATGAGAGCATGA 2267
QY 241 GAAGCACCAGAGATGCTGGTGGACCTCCAGTACAGCTGGGCCAAATCTTATGCCACAC 300
DB 2268 GAAGCACCAGAGATGCTGGTGGACCTCCAGTACAGCTGGGCCAAATCTTATGCCACAC 2327
QY 301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGATCATGTCAAAAATGGCGA 360
DB 2328 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGATCATGTCAAAAATGGCGA 2387
QY 361 TCTCTCAGAGCAGCATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGATATCTCAC 420
DB 2388 TCTCTCAGAGCAGCATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGATATCTCAC 2447
QY 421 ACGGAAGGCGTGTATGACAGGATGACCGGCTTCAGGGTCAATATCCCAAAATCGA 480

DB 2448 ACGGAAGGCGTGTATGACAGGATGACCGGCTTCAGGGTCAATATCCCAAAATCGA 2507
QY 481 CGAGAGCGCTTCATGATGAGA 501
DB 2508 CGAGAGCGCTTCATGATGAGA 2528

RESULT 3

ABK84968

ID ABK84968 standard; cDNA: 3705 BP.

XX ABK84968;

XX 13-AUG-2002 (first entry)

DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #3.

XX Human; autoimmune disease; haematopoietic disorder; DiGeorge syndrome;

KW blood protein disorder; agammaglobulinaemia; dysgammaaglobulinaemia;

KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;

KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;

KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;

KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;

KW autoimmune pulmonary inflammation; organ rejection; inflammation;

KW CLASP; gene; ss.

XX Homo sapiens.

OS WO200231117-A2.

XX 18-APR-2002.

XX 15-OCT-2001; 2001WO-US32202.

XX 13-OCT-2000; 2000US-0687837.

XX (ARBO-) ARBOR VITA CORP.

PA (GARW/) GARMAN J D.

PA (CAND/) CANDIA A F.

XX Lu PS;

XX WPI: 2002-416861/44.

DR P-PSDB: ABG61674.

XX New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating

PT an immune response, and for treating multiple sclerosis, rheumatoid

PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,

PT and sepsis

XX Disclosure: Figure 3A; 245pp; English.

XX The invention relates to an isolated polypeptide (I) comprising an amino

XX acid sequence that has 90 % sequence identity to one of the human

XX cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2a, 2b, 2C, 2E)

XX sequences (PS). (I) is useful for identifying a compound or agent that

XX binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for

XX detecting a CLASP-2 polypeptide in a sample. (II) is useful for

XX inhibiting an immune response in a subject. A pharmaceutical composition

XX comprising a nucleic acid encoding (I), or (II) is useful for preventing

XX or treating a CLASP-2 mediated disease e.g., an autoimmune disease, where

XX the autoimmune disease is caused or exacerbated by increased activity

XX of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for

XX detection or inhibition of CLASP-2 expression (e.g., antisense or

XX ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2

XX polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-

XX antibodies or are used as therapeutic polypeptides. The CLASP-2

XX polynucleotide or fragments can be used in diagnostics (e.g., as probes

XX for CLASP-2 expression), as a lymphocyte marker and for therapeutic

XX purposes. CLASP-2 polynucleotides can construct transgenic and knockout

XX animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2

XX polynucleotides can screen for CLASP-2 agonists and antagonists.

XX CLASP-2 polypeptides or polynucleotides can treat deficiencies or

disorders of the immune system, by activating or inhibiting the activation, differentiation of immune cells and can treat or detect deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides or polynucleotides can increase differentiation and proliferation of haematopoietic cells, including the pluripotent stem cells to treat those disorders associated with a decrease in certain (or many) types of haematopoietic cells e.g., immunologic deficiency syndromes including blood protein disorders (e.g., agammaglobulinaemia, common variable dysgamma globulinaemia, ataxia telangiectasia, thrombocytopenia, or immunodeficiency, DiGeorge syndrome, lymphopenia, or haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or detect autoimmune diseases, e.g., Addison's disease, hemolytic anaemia, Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus, endometriosis, autoimmune thyroiditis, and autoimmune pulmonary inflammation. CLASP-2 can be used to treat anaphylaxis or hypersensitivity to an antigenic molecule, organ rejection or graft-versus-host disease (GVHD) and inflammation. ABRK4922-ABR85018 represent cadherin-like asymmetry protein (CLASP) coding sequences and PCR primers of the invention.

Sequence 3705 BP: 1017 A; 911 C; 899 G; 878 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 3705;

Best Local Similarity 100.0%; Pred. No. 1e-141; Mismatches 0; Gaps 0;

Matches 501; Conservative 0; Indels 0;

1 TTACACTGGAAGAAGTCTTTGTCGGACACATTTCGATCATATCTGTACGCA 60

2022 TTACACTGGAAGAAGTCTTTGTCGGACACATTTCGATCATATCTGTACGCA 2081

61 GCTGATACAGACGTTGTTGGCATTGGGAAACAGATTCAGACAGTCCCTGCCATCAT 120

2082 GCTGATACAGACGTTGTTGGCATTGGGAAACAGATTCAGACAGTCCCTGCCATCAT 2141

121 CACACACTGTCCACACTGACCGGCTTATTAAGACACACAGCTTCCTGATGAGAA 180

2142 CACACACTGTCCACACTGACCGGCTTATTAAGACACACAGCTTCCTGATGAGAA 2201

181 GGACTTAAACCAAAAGATATGACGACGGTCTAATGGCCACCCCGATGAAGAGCATGA 240

2202 GGACTTAAACCAAAAGATATGACGACGGTCTAATGGCCACCCCGATGAAGAGCATGA 2261

241 GAAGCACCAGAGATGCTGTGGACCTCCAGTACAGCCTGGCCAAATCTTATGCCAGCAC 300

2262 GAAGCACCAGAGATGCTGTGGACCTCCAGTACAGCCTGGCCAAATCTTATGCCAGCAC 2321

301 GCCCAGCTCAGAGAGACGTGGCTGACAGCATGGCCAGATCCATGTCACAAATGGCGCA 360

2322 GCCCAGCTCAGAGAGACGTGGCTGACAGCATGGCCAGATCCATGTCACAAATGGCGCA 2381

361 TCTCTCAGAGGAGCATGCTATGTCACAGTACAGCCCTAGTGGCAGAAATCTCAGC 420

2382 TCTCTCAGAGGAGCATGCTATGTCACAGTACAGCCCTAGTGGCAGAAATCTCAGC 2441

421 ACGGAAGGCGCTTTTACAGCAAGATGACCCCTTCAGGGGTCAATTACCCCAACATCGA 480

2442 ACGGAAGGCGCTTTTACAGCAAGATGACCCCTTCAGGGGTCAATTACCCCAACATCGA 2501

481 CGAGGAGGCGCTTCATGATGGA 501

2502 CGAGGAGGCGCTTCATGATGGA 2522

RESULT 4
ABR84964
ID ABR84964 standard; cDNA; 4806 BP.

AC ABR84964;

DT 13-AUG-2002 (first entry)

DE DNA encoding cadherin-like asymmetry protein (CLASP).

Human; autoimmune disease; haematopoietic disorder; DiGeorge syndrome; ataxia telangiectasia; agammaglobulinaemia; dysgamma globulinaemia; thrombocytopenia; Addison's disease; Grave's disease; hemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus; endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity; autoimmune pulmonary inflammation; organ rejection; inflammation; CLASP; gene; ss.

Homo sapiens.

MO200231117-A2.

18-APR-2002.

15-OCT-2001; 2001WO-US32202.

13-OCT-2000; 2000US-0687837.

(ARB0-) ARBOR VITA CORP.

(GARM/) GARRAN J D.

(CAND/) CANDIA A F.

Lu PS;

WPI; 2002-416861/44.

P-PSDB; ABR61670.

New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating an immune response, and for treating multiple sclerosis, rheumatoid arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock, and sepsis.

Disclosure; Figure 1; 245pp; English.

The invention relates to an isolated polypeptide (I) comprising an amino acid sequence that has 90 % sequence identity to one of the human cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E) sequences (PS). (I) is useful for identifying a compound or agent that binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for detecting a CLASP-2 polypeptide in a sample. (II) is useful for inhibiting an immune response in a subject. A pharmaceutical composition comprising a nucleic acid encoding (I), or (II) is useful for preventing or treating a CLASP-2 mediated disease e.g., an autoimmune disease, where the autoimmune disease is caused or exacerbated by increased activity of rh1 cells. CLASP-2 polynucleotides are useful as probes or primers for detection or inhibition of CLASP-2 expression (e.g., antisense or ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2 polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-2 antibodies or are used as therapeutic polypeptides. The CLASP-2 polynucleotide or fragments can be used in diagnostics (e.g., as probes for CLASP-2 expression), as a lymphocyte marker and for therapeutic purposes. CLASP-2 polynucleotides can construct transgenic and knockout animals, e.g., for screening of CLASP-2 agonists and antagonists. CC polynucleotides can screen for CLASP-2 agonists and antagonists. CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or disorders of the immune system, by activating or inhibiting the activation, differentiation of immune cells and can treat or detect deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides or polynucleotides can increase differentiation and proliferation of haematopoietic cells, including the pluripotent stem cells to treat those disorders associated with a decrease in certain (or many) types of haematopoietic cells e.g., immunologic deficiency syndromes including blood protein disorders (e.g., agammaglobulinaemia, common variable dysgamma globulinaemia, ataxia telangiectasia, thrombocytopenia, or immunodeficiency, DiGeorge syndrome, lymphopenia, or haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or detect autoimmune diseases, e.g., Addison's disease, hemolytic anaemia, Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus, endometriosis, autoimmune thyroiditis, and autoimmune pulmonary inflammation. CLASP-2 can be used to treat anaphylaxis or hypersensitivity to an antigenic molecule, organ rejection or graft-versus-host disease (GVHD) and inflammation. ABRK4922-ABR85018

CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.

Sequence 4806 BP; 1352 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4806;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTACACTGGAAAGAGTCTTGTCCGGACACATTGCAAGTCATCATCTGTGACCA 60
Db 2499 TTACACTGGAAAGAGTCTTGTCCGGACACATTGCAAGTCATCATCTGTGACCA 2558
QY 61 GCTGATAGCAGAGCTGTGTCATTGGGAAACAGATTCACAGACTCCCTGTCATCAT 120
Db 2559 GCTGATAGCAGAGCTGTGTCATTGGGAAACAGATTCACAGACTCCCTGTCATCAT 2618
QY 121 CAACAACGTGTCACACAGTACCGGCTTATTTAAGCACACACAGCTTCTCTGTGATGAA 180
Db 2619 CAACAACGTGTCACACAGTACCGGCTTATTTAAGCACACACAGCTTCTCTGTGATGAA 2678
QY 181 GGACTTAACCAAAAGATACGACGCTGCTAATGGCCACCGCCAGATGAAGAGCATGA 240
Db 2679 GGACTTAACCAAAAGATACGACGCTGCTAATGGCCACCGCCAGATGAAGAGCATGA 2738
QY 241 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGCTGCGCAATCCTATGCCAGCAC 300
Db 2739 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGCTGCGCAATCCTATGCCAGCAC 2798
QY 301 GCCCGAGCTCAGAACAGCTGCTGCACAGCATGCGCAGATTCATGCAAAATGGCGA 360
Db 2799 GCCCGAGCTCAGAACAGCTGCTGCACAGCATGCGCAGATTCATGCAAAATGGCGA 2858
QY 361 TCTCTAGAGAGCGCATGTGCTATGCCAGTACAGCGCTAGTGGCAATATCTGCAC 420
Db 2859 TCTCTAGAGAGCGCATGTGCTATGCCAGTACAGCGCTAGTGGCAATATCTGCAC 2918
QY 421 ACGGAAAGCGGTGTTAGACAAAGATGACACCGCTTCAGGGTCAATACCCCAACATGCA 480
Db 2919 ACGGAAAGCGGTGTTAGACAAAGATGACACCGCTTCAGGGTCAATACCCCAACATGCA 2978
QY 481 CGAGGAGGCTCCATGATGCA 501
Db 2979 CGAGGAGGCTCCATGATGCA 2999
```

RESULT 5
AAC87972
ID AAC87972 standard; cDNA: 4807 BP.
XX
AC AAC87972;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CLASP-2 nucleotide sequence.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antirheumatic; cytoskeletal;
KW hypotensive; antihemetic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.

```
XX 14-APR-1999; 99US-0129171.  
PR 14-MAY-1999; 99US-0134114.  
PR 14-MAY-1999; 99US-0134117.  
PR 14-MAY-1999; 99US-0134118.  
PR 21-OCT-1999; 99US-0160860.  
PR 29-OCT-1999; 99US-0162498.  
PR 13-DEC-1999; 99US-0170453.  
PR 14-JAN-2000; 99US-0176195.  
PR 14-FEB-2000; 2000US-0182296.  
XX  
XX (ARBO-) ARBOR VITA CORP.  
XX  
XX Lu PS;  
XX WPI: 2000-619230/59.  
XX P-PSDB: AAB36527.  
XX  
XX Isolated cadherin-like asymmetry protein-2 polynucleotide and  
XX polypeptide used to diagnose, treat and prevent autoimmune diseases and  
XX inflammatory responses -  
XX  
XX Example 1; Fig 1; 286bp; English.  
XX  
XX The present invention describes cadherin-like asymmetry protein-2  
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,  
XX antiinflammatory, antirheumatic, cytoskeletal, hypotensive, antirheumatic,  
XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be  
XX used to inhibit an immune response in a subject by interfering with the  
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An  
XX immune response in a subject may also be inhibited by administering an  
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,  
XX proteins and antibodies can be used to prevent or treat a CLASP-2  
XX mediated disease, such as an autoimmune disease caused or exacerbated  
XX by increased activity of TH1 cells. They can also be used to treat  
XX hypersensitivities, prevent transplantation rejection responses and  
XX augment immune responsiveness in immunodeficiency states, inhibit  
XX proliferation and differentiation of cells involved in an inflammatory  
XX response e.g. arthritis, inflammatory bowel disease and increase  
XX differentiation and proliferation of haematopoietic cells e.g. to treat  
XX anaemia, thrombocytopaenia and other blood protein disorders. Disorders  
XX treated by disrupting CLASP-2 function include multiple sclerosis,  
XX rheumatoid arthritis, endometriosis and pregnancy induced hypertension.  
XX The present sequence encodes human CLASP-2, which is used in the  
XX exemplification of the present invention.  
XX  
XX Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;  
SQ
```

Query Match 100.0%; Score 501; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTACACTGGAAAGAGTCTTGTCCGGACACATTGCAAGTCATCATCTGTGACCA 60
Db 2500 TTACACTGGAAAGAGTCTTGTCCGGACACATTGCAAGTCATCATCTGTGACCA 2559
QY 61 GCTGATAGCAGAGCTGTGTCATTGGGAAACAGATTCACAGACTCCCTGTCATCAT 120
Db 2560 GCTGATAGCAGAGCTGTGTCATTGGGAAACAGATTCACAGACTCCCTGTCATCAT 2619
QY 121 CAACAACGTGTCACACAGTACCGGCTTATTTAAGCACACAGCTTCTCTGTGATGAA 180
Db 2620 CAACAACGTGTCACACAGTACCGGCTTATTTAAGCACACAGCTTCTCTGTGATGAA 2679
QY 181 GGACTTAACCAAAAGATACGACGCTGCTAATGGCCACCGCCAGATGAAGAGCATGA 240
Db 2680 GGACTTAACCAAAAGATACGACGCTGCTAATGGCCACCGCCAGATGAAGAGCATGA 2739
QY 241 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGCTGCGCAATCCTATGCCAGCAC 300
Db 2740 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGCTGCGCAATCCTATGCCAGCAC 2799
QY 301 GCCCGAGCTCAGAACAGTGTGCTGCACAGCATGCGCAGATTCATGCAAAATGGCGA 360
```

|||||
DB 2800 GCCCGAGCTCAGGAAGAGCTGGCTGCAGACGATGGCCAGGATCCATGTCACAAAGTGGCA 2859
QY 361 TCTCTCAGAGGAGCAGCAATGCTGTATGTCACGTAACAGCCCTAGTGCGAGAATATCTCAC 420
DB 2860 TCTCTCAGAGGAGCAGCAATGCTGTATGTCACGTAACAGCCCTAGTGCGAGAATATCTCAC 2919
QY 421 ACGGAAGGCGGTGTAGACAGGATGCACCGCTTCAGGGTCATTACCCCAACATCGA 480
DB 2920 ACGGAAGGCGGTGTAGACAGGATGCACCGCTTCAGGGTCATTACCCCAACATCGA 2979
QY 481 CGAGGAGGCGCTCCATGATCGA 501
DB 2980 CGAGGAGGCGCTCCATGATCGA 3000
RESULT 6
AAC87973
ID AAC87973 standard; cDNA; 4807 BP.
XX
AC AAC87973;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CLASP-2A nucleotide sequence.
XX
KW CLASP-1: CLASP-2: transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antihistaminic; antianemic; haemostatic; neuroprotective;
KW hypotensivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
KW endometritis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO20061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI: 2000-619230/59.
DR P-PSDB; AAB36528.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PI inflammatory responses -
XX
PS Example 1; Fig 2B; 286bp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antihistaminic, cytostatic, hypotensive, antirheumatic,
CC antianemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An

CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence encodes human CLASP-2A, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;
Query Match 100.0%; Score 501; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 1,1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACACTGGAAGAAGTCTTTGTCGAGACATTTGCAAGTCATATCTGTACGCCA 60
DB 2500 TTACACTGGAAGAAGTCTTTGTCGAGACATTTGCAAGTCATATCTGTACGCCA 2559
QY 61 GCTATATGACAGCTTTTGGCATTTGGGGAACCAGATTCCAGACGTCCCTGTCCATCAT 120
DB 2560 GCTATATGACAGCTTTTGGCATTTGGGGAACCAGATTCCAGACGTCCCTGTCCATCAT 2619
QY 121 CAACACGTGTCACACAGTGCAGCGCTTTATTAAGCACACACAGCTTCCTGTATGTGAA 180
DB 2620 CAACACGTGTCACACAGTGCAGCGCTTTATTAAGCACACACAGCTTCCTGTATGTGAA 2679
QY 181 GGACTTAACCAAAAGATAGCACAGTGTATGTGCGCCAGCCCAAGTAGAGAGATGA 240
DB 2680 GGACTTAACCAAAAGATAGCACAGTGTATGTGCGCCAGCCCAAGTAGAGAGATGA 2739
QY 241 GAACGACCCAGAGTGTGTGTGACCTTCACAGTACAGCTTGCCCAATTCATGCCAGCAC 300
DB 2740 GAACGACCCAGAGTGTGTGTGACCTTCACAGTACAGCTTGCCCAATTCATGCCAGCAC 300
QY 301 GCCGAGCTCAGGAAGAGCTGGCTGCAGACATGAGCCAGATCCATGTCAAAATGGCGCA 360
DB 2800 GCCGAGCTCAGGAAGAGCTGGCTGCAGACATGAGCCAGATCCATGTCAAAATGGCGCA 2859
QY 361 TCTCTCAGAGGAGCAGCAATGCTGTATGTCACGTAACAGCCCTAGTGCGAGAATATCTCAC 420
DB 2860 TCTCTCAGAGGAGCAGCAATGCTGTATGTCACGTAACAGCCCTAGTGCGAGAATATCTCAC 2919
QY 421 ACGGAAGGCGGTGTAGACAGGATGCACCGCTTCAGGGTCATTACCCCAACATCGA 480
DB 2920 ACGGAAGGCGGTGTAGACAGGATGCACCGCTTCAGGGTCATTACCCCAACATCGA 2979
QY 481 CGAGGAGGCGCTCCATGATCGA 501
DB 2980 CGAGGAGGCGCTCCATGATCGA 3000
RESULT 7
ABK84966
ID ABK84966 standard; cDNA; 4807 BP.
XX
AC ABK84966;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #1.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digesorge syndrome;
KW blood protein disorder; agammaglobulinemia; dysgammaglobulinemia;
KW ataxia telangiectasia; common variable immunodeficiency; Lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;

KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
 KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
 KW autoimmune pulmonary inflammation; organ rejection; inflammation;
 KW CLASP; gene: ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200231117-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-US32202.
 XX
 PR 13-OCT-2000; 2000US-0687837.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 PA (GARW/) GARMAN J D.
 PA (CAND/) CANDIA A F.
 XX
 PI Lu PS:
 XX
 DR WPI; 2002-416861/44.
 DR P-PSDB; AB661672.
 XX
 PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
 PT an immune response, and for treating multiple sclerosis, rheumatoid
 PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
 PT and sepsis
 PT
 PS Disclosure: Figure 3A: 245bp; English.
 XX
 CC The invention relates to an isolated polypeptide (I) comprising an amino
 CC acid sequence that has 90 % sequence identity to one of the human
 CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2a, 2b, 2C, 2E)
 CC sequences (PS). (I) is useful for identifying a compound or agent that
 CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
 CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
 CC inhibiting a immune response in a subject. A pharmaceutical composition
 CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
 CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
 CC the autoimmune disease is caused or exacerbated by increased activity
 CC of TML cells. CLASP-2 polynucleotides are useful as probes or primers for
 CC detection or inhibition of CLASP-2 expression (e.g., antisense or
 CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
 CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
 CC antibodies or are used as therapeutic polypeptides. The CLASP-2
 CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
 CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
 CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
 CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
 CC polynucleotides can screen for CLASP-2 agonists and antagonists.
 CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
 CC disorders of the immune system, by activating or inhibiting the
 CC activation, differentiation of immune cells and can treat or detect
 CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
 CC or polynucleotides can increase differentiation and proliferation of
 CC haematopoietic cells, including the pluripotent stem cells to treat those
 CC disorders associated with a decrease in certain (or many) types of
 CC haematopoietic cells e.g., immunologic deficiency syndromes including
 CC blood protein disorders (e.g., agammaglobulinemia,
 CC dysgammaglobulinemia, ataxia telangiectasia, common variable
 CC immunodeficiency, DiGeorge syndrome, lymphopenia, thrombocytopenia, or
 CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
 CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
 CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
 CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
 CC inflammation. CLASP-2 can be used to treat anaphylaxis or
 CC hypersensitivity to an antigenic molecule, organ rejection or graft-
 CC versus-host disease (GVHD) and inflammation. ABK4922-ABK85018
 CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
 CC PCR primers of the invention.
 XX
 SO Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4807;
 Best Local Similarity 100.0%; Pred. No. 1,1e-141;
 Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACACTGGAAGAAGTCCCTTTGTCGGACACATTTGCAAGTCATCATCTGTCAACCA 60
 DB 2500 TTACACTGGAAGAAGTCCCTTTGTCGGACACATTTGCAAGTCATCATCTGTCAACCA 2559
 QY 61 GCTGATGACGACGACGCTGTTGGCAATGGGAACCGATATTCAGCAGTCCGTGCATCAT 120
 DB 2560 GCTGATGACGACGACGCTGTTGGCAATGGGAACCGATATTCAGCAGTCCGTGCATCAT 2619
 QY 121 CAACACTGTGCCACACAGTACCGCGCTATTATAGCACACAGCTTCTCTCTGATGGAA 180
 DB 2620 CAACACTGTGCCACACAGTACCGCGCTATTATAGCACACAGCTTCTCTCTGATGGAA 2679
 QY 181 GGACTTAACCAAAAGATACGACGCGTGTATGAGCCACCGCCAGATGAAGAGCATGA 240
 DB 2680 GGACTTAACCAAAAGATACGACGCGTGTATGAGCCACCGCCAGATGAAGAGCATGA 2739
 QY 241 GAACGACCCAGAGATGCTGTGTGACCTCCAGTACAGCTGTGCCAAATCTTATGCGACAC 300
 DB 2740 GAACGACCCAGAGATGCTGTGTGACCTCCAGTACAGCTGTGCCAAATCTTATGCGACAC 2799
 QY 301 GCCCGAGCTCAGGACGACGCGTGTGACGACATGGCCAGATGCCAATGCAAAATGGCGCA 360
 DB 2800 GCCCGAGCTCAGGACGACGCGTGTGACGACATGGCCAGATGCCAATGCAAAATGGCGCA 2859
 QY 361 TCTCTCAGAGGACGACCAATGTGTATGTCCAGCTAACAGCCCTAGTGCGCAATATCTCAC 420
 DB 2860 TCTCTCAGAGGACGACCAATGTGTATGTCCAGCTAACAGCCCTAGTGCGCAATATCTCAC 2919
 QY 421 ACGGAAGGGGTGTTATGACCAAGGATGACGCGCTTCAGGGGTCATTCGCCAAACATCGA 480
 DB 2920 ACGGAAGGGGTGTTATGACCAAGGATGACGCGCTTCAGGGGTCATTCGCCAAACATCGA 2979
 QY 481 CGAAGAGCGCTTCATGATGGA 501
 DB 2980 CGAAGAGCGCTTCATGATGGA 3000

RESULT 8
 ABK84973
 ID ABK84973 standard; DNA; 4807 BP.
 XX
 AC ABK84973;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform.
 XX
 KW Human; autoimmune disease; haematopoietic disorder; DiGeorge syndrome;
 KW blood protein disorder; agammaglobulinemia; dysgammaglobulinemia;
 KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
 KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
 KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
 KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
 KW autoimmune pulmonary inflammation; organ rejection; inflammation;
 KW CLASP; gene: ds.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO200231117-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-US32202.
 XX
 PR 13-OCT-2000; 2000US-0687837.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 PA (GARW/) GARMAN J D.

PA (CAND/) CANDIA A F.
 XX
 XX
 PI Lu PS:
 XX
 DR WPI: 2002-416861/44.
 DR P-PSDB: ABG61686.
 XX
 PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
 PT an immune response, and for treating sclerosis, rheumatoid
 PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
 PT and sepsis -
 XX
 XX Example 4: Figure 6A: 245bp: English.
 PS
 SS
 CC The invention relates to an isolated polypeptide (I) comprising an amino
 CC acid sequence that has 90 % sequence identity to one of the human
 CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
 CC sequences (PS). (I) is useful for identifying a compound or agent that
 CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
 CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
 CC inhibiting a immune response in a subject. A pharmaceutical composition
 CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
 CC or treating a CLASP-2 mediated disease e.g., an autoimmune disease, where
 CC the autoimmune disease is caused or exacerbated by increased activity
 CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
 CC detection or inhibition of CLASP-2 expression (e.g., antisense or
 CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
 CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
 CC antibodies or are used as therapeutic polypeptides. The CLASP-2
 CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
 CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
 CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
 CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
 CC polynucleotides can screen for CLASP-2 agonists and antagonists.
 CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
 CC disorders of the immune system, by activating or inhibiting the
 CC activation, differentiation of immune cells and can treat or detect
 CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
 CC or polynucleotides can increase differentiation and proliferation of
 CC haematopoietic cells, including the pluripotent stem cells to treat those
 CC disorders associated with a decrease in certain (or many) types of
 CC haematopoietic cells e.g., immunologic deficiency syndromes including
 CC blood protein disorders (e.g., agammaglobulinemia,
 CC dysgamma globulinemia, ataxia telangiectasia, common variable
 CC immunodeficiency, DiGeorge syndrome, lymphopenia, thrombocytopenia, or
 CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
 CC detect autoimmune diseases, e.g., Addison's disease, hemolytic anaemia,
 CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
 CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
 CC inflammation. CLASP-2 can be used to treat anaphylaxis or
 CC hypersensitivity to an antigenic molecules, organ rejection or graft-
 CC versus-host disease (GVHD) and inflammation. ABRK4922-ABK85018
 CC represent cadherin-like asymmetry protein (CLASP) coding sequences
 CC and PCR primers of the invention.
 CC
 XX
 XX Sequence 4807 BP: 1353 A; 1117 C; 1136 G; 1201 T; 0 other:
 SO
 Query Match 100.0%; Score 501; DB 24; Length 4807;
 Best Local Similarity 100.0%; Pred. NO. 1,1e-141;
 Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTACACTGGAAGAAGTCTTGTCCGACACATTGGACATTCATATCTGTACGCA 60
 DB 2500 TTACACTGGAAGAAGTCTTGTCCGACACATTGGACATTCATATCTGTACGCA 2559
 QY 61 GGTGATAGCAGACGTTGTCGATTGGGGAACCAAGATTCCAGAGCTCCCTGCATCAT 120
 DB 2560 GGTGATAGCAGACGTTGTCGATTGGGGAACCAAGATTCCAGAGCTCCCTGCATCAT 2619
 QY 121 CACACACTGTGCGACAGAGCGGCTTATTAAGCAGACACACTTCTCCTCGATGTGAA 180
 DB 2620 CACACACTGTGCGACAGAGCGGCTTATTAAGCAGACACACTTCTCCTCGATGTGAA 2679

QY 181 GGACTTAACCAAAAGATACGACGAGTGTCTATGTGGCCACCGCCGAGATGAGAGCATGA 240
 DB 2680 GGACTTAACCAAAAGATACGACGAGTGTCTATGTGGCCACCGCCGAGATGAGAGCATGA 2739
 QY 241 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGCCTGGCCAAATCTTATGCCAGCAC 300
 DB 2740 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGCCTGGCCAAATCTTATGCCAGCAC 2799
 QY 301 GCCCGAGCTCAGACAGCTGCTGCACAGCATTGGCCAGATTCATGTCCAAAATGGCGGA 360
 DB 2800 GCCCGAGCTCAGACAGCTGCTGCACAGCATTGGCCAGATTCATGTCCAAAATGGCGGA 2859
 QY 361 TCTCTCAGAGCAGCATGTGCTATGCCATACAGCCTAGTGGAGATATCTCAC 420
 DB 2860 TCTCTCAGAGCAGCATGTGCTATGCCATACAGCCTAGTGGAGATATCTCAC 2919
 QY 421 ACGGAAGGCGGTGTTAGACAGAGATGACACCGCTTCAGGCTCATTTACCCCAACATGGA 480
 DB 2920 ACGGAAGGCGGTGTTAGACAGAGATGACACCGCTTCAGGCTCATTTACCCCAACATGGA 2979
 QY 481 CGAGAGGCTCCATCATGGA 501
 DB 2980 CGAGAGGCTCCATCATGGA 3000
 RESULT 9
 AAC87974
 ID AAC87974 standard; cDNA; 4898 BP.
 XX
 XX AAC87974:
 AC 07-MAR-2001 (first entry)
 DT
 XX Preliminary CLASP-2 nucleotide sequence #1.
 DE
 XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
 KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
 KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
 KW hypotensive; antirheumatic; antinaeemic; haemostatic; neuroprotective;
 KW hypersensitivity; transplantation rejection response; immunodeficiency;
 KW proliferation; differentiation; inflammatory response; arthritis;
 KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
 KW anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
 KW endometriosis; pregnancy induced hypertension; ss.
 KW
 OS Homo sapiens.
 XX
 XX WO200061747-A2.
 PN 19-OCT-2000.
 PD
 XX 13-APR-2000; 2000WO-US10158.
 PF
 XX 14-APR-1999; 99US-0129171.
 PR 14-MAY-1999; 99US-0134114.
 PR 14-MAY-1999; 99US-0134117.
 PR 14-MAY-1999; 99US-0134118.
 PR 21-OCT-1999; 99US-0160860.
 PR 29-OCT-1999; 99US-0162498.
 PR 13-DEC-1999; 99US-0170453.
 PR 14-JAN-2000; 2000US-0176195.
 PR 14-FEB-2000; 2000US-0182296.
 XX
 XX (ARBO-) ARBOR VITA CORP.
 XX
 XX Lu PS:
 PI WPI: 2000-619230/59.
 DR P-PSDB: AAB36529.
 XX
 PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
 PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
 PT inflammatory responses -


```
QY 1 TTACACTGGAAGAAGTCTTTGTCCGGACACATTTGCAATGATATATCTGTACGCCA 60
   |||||||
Db 2591 TTACACTGGAAGAAGTCTTTGTCCGGACACATTTGCAATGATATATCTGTACGCCA 2650
QY 61 GCTGATAGCAGACGTTTGTGGCATTTGGGGAACACCATTTCCACAGTCCCTGTCCATCAT 120
   |||||||
Db 2651 GCTGATAGCAGACGTTTGTGGCATTTGGGGAACACCATTTCCACAGTCCCTGTCCATCAT 2710
QY 121 CAACACTGTGGCCAAAGTGCACGGCTTATTAAAGCACACACAGCTTCTCTGTATGTGAA 180
   |||||||
Db 2711 CAACACTGTGGCCAAAGTGCACGGCTTATTAAAGCACACACAGCTTCTCTGTATGTGAA 2770
QY 181 GGACTTAACCAAAAAGATATACGACAGTGTCTAATATGCCCCAGCCAGATGAAGACATGA 240
   |||||||
Db 2771 GGACTTAACCAAAAAGATATACGACAGTGTCTAATATGCCCCAGCCAGATGAAGACATGA 2830
QY 241 GAACGACCCAGAGATGCTGTGTGACCTTCACGTACAGCTTGGCCAAATCTTATGCCAGCAC 300
   |||||||
Db 2831 GAACGACCCAGAGATGCTGTGTGACCTTCACGTACAGCTTGGCCAAATCTTATGCCAGCAC 2890
QY 301 GCCCGAGCTCGAAGAAGCTGGCTCGACAGATGGCCAGATCCATGTCAAAAATGCGCA 360
   |||||||
Db 2891 GCCCGAGCTCGAAGAAGCTGGCTCGACAGATGGCCAGATCCATGTCAAAAATGCGCA 2950
QY 361 TCTCTCAGAGGCACCAATGTGCTATGTCCACAGTAAACGCCCTAGTGCAGAAATATCTCAC 420
   |||||||
Db 2951 TCTCTCAGAGGCACCAATGTGCTATGTCCACAGTAAACGCCCTAGTGCAGAAATATCTCAC 3010
QY 421 ACGGAAGGCGTGTATTAGACAAAGATGCACCGCTTACAGGCTCATTAACCCCAACATCGA 480
   |||||||
Db 3011 ACGGAAGGCGTGTATTAGACAAAGATGCACCGCTTACAGGCTCATTAACCCCAACATCGA 3070
QY 481 CGAGGAGGCTTCATGATGA 501
   |||||||
Db 3071 CGAGGAGGCTTCATGATGA 3091

RESULT 11
AAC87976
ID AAC87976 standard; cDNA; 4898 BP.
XX
AC AAC87976;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #3.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000MO-US10158.
XX
PR 14-APR-1999; 9905-0129171.
PR 14-MAY-1999; 9905-0134114.
PR 14-MAY-1999; 9905-0134117.
PR 14-MAY-1999; 9905-0134118.
PR 21-OCT-1999; 9905-0160860.
PR 29-OCT-1999; 9905-0162498.
PR 13-DEC-1999; 9905-0170453.
PR 14-JAN-2000; 2000US-0176195.
```

```
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS:
XX WPI: 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure: Fig 10C; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states. Inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 1,1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACACTGGAAGAAGTCTTTGTCCGGACACATTTGCAATGATATATCTGTACGCCA 60
   |||||||
Db 2591 TTACACTGGAAGAAGTCTTTGTCCGGACACATTTGCAATGATATATCTGTACGCCA 2650
QY 61 GCTGATAGCAGACGTTTGTGGCATTTGGGGAACACCATTTCCACAGTCCCTGTCCATCAT 120
   |||||||
Db 2651 GCTGATAGCAGACGTTTGTGGCATTTGGGGAACACCATTTCCACAGTCCCTGTCCATCAT 2710
QY 121 CAACACTGTGGCCAAAGTGCACGGCTTATTAAAGCACACACAGTCTCTCTGTATGTGAA 180
   |||||||
Db 2711 CAACACTGTGGCCAAAGTGCACGGCTTATTAAAGCACACACAGTCTCTCTGTATGTGAA 2770
QY 181 GGACTTAACCAAAAAGATATACGACAGTGTCTAATATGCCCCAGCCAGATGAAGACATGA 240
   |||||||
Db 2771 GGACTTAACCAAAAAGATATACGACAGTGTCTAATATGCCCCAGCCAGATGAAGACATGA 2830
QY 241 GAACGACCCAGAGATGCTGTGTGACCTTCACGTACAGCTTGGCCAAATCTTATGCCAGCAC 300
   |||||||
Db 2831 GAACGACCCAGAGATGCTGTGTGACCTTCACGTACAGCTTGGCCAAATCTTATGCCAGCAC 2890
QY 301 GCCCGAGCTCGAAGAAGCTGGCTCGACAGATGGCCAGATCCATGTCAAAAATGCGCA 360
   |||||||
Db 2891 GCCCGAGCTCGAAGAAGCTGGCTCGACAGATGGCCAGATCCATGTCAAAAATGCGCA 2950
QY 361 TCTCTCAGAGGCACCAATGTGCTATGTCCACAGTAAACGCCCTAGTGCAGAAATATCTCAC 420
   |||||||
Db 2951 TCTCTCAGAGGCACCAATGTGCTATGTCCACAGTAAACGCCCTAGTGCAGAAATATCTCAC 3010
QY 421 ACGGAAGGCGTGTATTAGACAAAGATGCACCGCTTACAGGCTCATTAACCCCAACATCGA 480
   |||||||
Db 3011 ACGGAAGGCGTGTATTAGACAAAGATGCACCGCTTACAGGCTCATTAACCCCAACATCGA 3070
```

OY 481 CGAGAGGCGCTCCATGATGGA 501
|||||
Db 3071 CGAGAGGCGCTCCATGATGGA 3091

RESULT 12

AAC87977

ID AAC87977 standard; cDNA; 4898 BP.

XX AAC87977;

XX 07-MAR-2001 (first entry)

XX Preliminary CLASP-2 nucleotide sequence #4.

XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KM cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KM immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KM hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KM hypersensitivity; transplantation rejection response; immunodeficiency;
KM proliferation; differentiation; inflammatory response; arthritis;
KM inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KM anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KM endometrios; pregnancy induced hypertension; ss.

XX Homo sapiens.

XX WO200061747-A2.

XX 19-OCT-2000.

XX 13-APR-2000; 2000MO-US10158.

XX 14-APR-1999; 99US-0129171.

XX 14-MAY-1999; 99US-0134114.

XX 14-MAY-1999; 99US-0134117.

XX 21-OCT-1999; 99US-0134118.

XX 29-OCT-1999; 99US-0162498.

XX 13-DEC-1999; 99US-0170453.

XX 14-JAN-2000; 2000US-0176195.

XX 14-FEB-2000; 2000US-0182296.

XX (ARBO-) ARBOR VITA CORP.

XX Lu PS;

XX MPI; 2000-619230/59.

XX Isolated cadherin-like asymmetry protein-2 polynucleotide and

XX polypeptide used to diagnose, treat and prevent autoimmune diseases and

XX inflammatory responses -

XX Disclosure; Fig 10D; 286pp; English.

XX The present invention describes cadherin-like asymmetry protein-2

XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,

XX antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,

XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be

XX used to inhibit an immune response in a subject by interfering with the

XX ability of a CLASP-2 protein to bind to another T cell or B cell. An

XX immune response in a subject may also be inhibited by administering an

XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,

XX proteins and antibodies can be used to prevent or treat a CLASP-2

XX mediated disease, such as an autoimmune disease caused or exacerbated

XX by increased activity of Th1 cells. They can also be used to treat

XX hypersensitivities, prevent transplantation rejection responses and

XX augment immune responsiveness in immunodeficiency states, inhibit

XX proliferation and differentiation of cells involved in an inflammatory

XX response e.g. arthritis, inflammatory bowel disease and increase

XX differentiation and proliferation of haematopoietic cells e.g. to treat

XX anaemia, thrombocytopaenia and other blood protein disorders. Disorders

CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometrios and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.

XX SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

XX Query Match 100.0%; Score 501; DB 21; Length 4898;

XX Best Local Similarity 100.0%; Pred. No. 1,1e-141;

XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTACACTGGAAGAAGTCTTTGTGTCGACACATTTGCAAGTCAATCATATGTCACCA 60
|||||
Db 2591 TTACACTGGAAGAAGTCTTTGTGTCGACACATTTGCAAGTCAATCATATGTCACCA 2650
61 GCTATAGCAGACGTTGTGTCATGCGGAAACAGATTCAGCAGTCCCTGTCATCAT 120
|||||
Db 2651 GCTATAGCAGACGTTGTGTCATGCGGAAACAGATTCAGCAGTCCCTGTCATCAT 2710
121 CAACAACGTGTCACACAGTGCACCGCTTATTAAGCACACAGCTTCTCTGATGTGAA 180
|||||
Db 2711 CAACAACGTGTCACACAGTGCACCGCTTATTAAGCACACAGCTTCTCTGATGTGAA 2770
181 GCACCTAACCAAAAGGATACGACGTCGTAATGCGCACCGCCAGATGAGAGCATGA 240
|||||
Db 2771 GCACCTAACCAAAAGGATACGACGTCGTAATGCGCACCGCCAGATGAGAGCATGA 2830
241 GAAGCAGCCAGAGATGCTGTGACCTCCAGTACAGCTGCGCAAAATCTATGCGACAC 300
|||||
Db 2831 GAAGCAGCCAGAGATGCTGTGACCTCCAGTACAGCTGCGCAAAATCTATGCGACAC 2890
301 GCCGAGCTCAGAGAGACGTGGCTGCACAGTGCSCAGATGCCATGTCAAAAATGGCGA 360
|||||
Db 2891 GCCGAGCTCAGAGAGACGTGGCTGCACAGTGCSCAGATGCCATGTCAAAAATGGCGA 2950
2951 TCTCTCAGAGGAGCAGCAATGTGCTATGTCACGTAACACCCCTAGGCGAGATATCTCAC 3010
361 TCTCTCAGAGGAGCAGCAATGTGCTATGTCACGTAACACCCCTAGGCGAGATATCTCAC 420
|||||
OY 421 ACGGAAGGCGGTGTAGACAGAGATGCACGCGCTTCAGGGGTCATTACCCAAACATCGA 480
3011 ACGGAAGGCGGTGTAGACAGAGATGCACGCGCTTCAGGGGTCATTACCCAAACATCGA 3070

OY 481 CGAGAGGCGCTCCATGATGGA 501
|||||
Db 3071 CGAGAGGCGCTCCATGATGGA 3091

RESULT 13

AAC87978

ID AAC87978 standard; cDNA; 4898 BP.

XX AAC87978;

XX 07-MAR-2001 (first entry)

XX Preliminary CLASP-2 nucleotide sequence #5.

XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KM cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KM immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KM hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KM hypersensitivity; transplantation rejection response; immunodeficiency;
KM proliferation; differentiation; inflammatory response; arthritis;
KM inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KM anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KM endometrios; pregnancy induced hypertension; ss.

XX Homo sapiens.

XX WO200061747-A2.

XX 19-OCT-2000.

```
XX 13-APR-2000; 2000WO-US10158.
PF
XX 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX Lu PS:
XX WPI: 2000-619230/59.
XX
XX Isolated cadherin-like asymmetry protein-2 polynucleotide and
XX polypeptide used to diagnose, treat and prevent autoimmune diseases and
XX inflammatory responses -
XX
XX Disclosure: Fig 10E; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antihyaluronic, cytostatic, hypotensive, antirheumatic,
XX antianemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX immune response in a subject may also be inhibited by administering an
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX proteins and antibodies can be used to prevent or treat a CLASP-2
XX mediated disease, such as an autoimmune disease caused or exacerbated
XX by increased activity of TH1 cells. They can also be used to treat
XX hypersensitivities, prevent transplantation rejection responses and
XX augment immune responsiveness in immunodeficiency states. Inhibit
XX proliferation and differentiation of cells involved in an inflammatory
XX response e.g. arthritis, inflammatory bowel disease and increase
XX differentiation and proliferation of haematopoietic cells e.g. to treat
XX anaemia, thrombocytopenia and other blood protein disorders. Disorders
XX treated by disrupting CLASP-2 function include multiple sclerosis,
XX rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
XX The present sequence represents a preliminary CLASP-2 nucleotide
XX sequence, from the present invention.
XX
XX Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other:
SQ
XX
XX Query Match 100.0%; Score 501; DB 21; Length 4898;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-141;
XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TTTCACCTGGAAAGAAAGTCCCTTTGTCGGACACATTTGGCAAGCATCATATTCGTACGCCA 60
XX |
XX Db 2591 TTACACCTGGAAAGAAAGTCCCTTTGTCGGACACATTTGGCAAGCATCATATTCGTACGCCA 2650
XX |
XX QY 61 GCTGATAGCAGAGCTTGTGGCATTTGGGAAACACAGATTCAGAGTCCCTGTCATCAT 120
XX |
XX Db 2651 GCTGATAGCAGAGCTTGTGGCATTTGGGAAACACAGATTCAGAGTCCCTGTCATCAT 2710
XX |
XX QY 121 CAACAACTGTGCCAAGAGTACCGGCTTATTAAGACACACAGCTTCTCTGTGATGTGAA 180
XX |
XX Db 2711 CAACAACTGTGCCAAGAGTACCGGCTTATTAAGACACACAGCTTCTCTGTGATGTGAA 2770
XX |
XX QY 181 GGACTTACCAAAAGATAGCAGCGTGTATGAGCGCAGCCGCAATATAGAGACATGA 240
XX |
XX Db 2771 GGACTTACCAAAAGATAGCAGCGTGTATGAGCGCAGCCGCAATATAGAGACATGA 2830
XX |
XX QY 241 GAACGAGCCAGAGATGTGTGTGAGACCTCCAGTACGCTGAGCCCAATCTATGCGCAGAC 300
XX |
XX Db 2831 GAACGAGCCAGAGATGTGTGTGAGACCTCCAGTACGCTGAGCCCAATCTATGCGCAGAC 2890
```

```
QY 301 GCCCGAGCTCAGAGACAGCTGCTCGACAGCATGCGCAGGATTCATGCAAAATGGCGA 360
XX |
XX Db 2891 GCCCGAGCTCAGAGACAGCTGCTCGACAGCATGCGCAGGATTCATGCAAAATGGCGA 2950
XX |
XX QY 361 TCTCTCAGAGGAGCAATGTGCTATGTCCAGTAACAGCCCTAGTGGCAATATCTCAC 420
XX |
XX Db 2951 TCTCTCAGAGGAGCAATGTGCTATGTCCAGTAACAGCCCTAGTGGCAATATCTCAC 3010
XX |
XX QY 421 ACGGAAAGGCGGTGTTAGACAGAGATGCACCGCTTCAGGGTCATTATCCCAACATGCA 480
XX |
XX Db 3011 ACGGAAAGGCGGTGTTAGACAGAGATGCACCGCTTCAGGGTCATTATCCCAACATGCA 3070
XX |
XX QY 481 CGAGAGAGCCCTCCATGATGCA 501
XX |
XX Db 3071 CGAGAGAGCCCTCCATGATGCA 3091
XX
XX RESULT 14
XX AAC87979
XX ID AAC87979 standard; cDNA; 4898 BP.
XX
XX AAC87979;
XX
XX 07-MAR-2001 (first entry)
XX
XX Preliminary CLASP-2 nucleotide sequence #6.
XX
XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
XX cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
XX immunomodulatory; antiinflammatory; antihyaluronic; cytostatic;
XX hypotensive; antirheumatic; antianemic; haemostatic; neuroprotective;
XX hypersensitivity; transplantation rejection response; immunodeficiency;
XX proliferation; differentiation; inflammatory response; arthritis;
XX inflammatory bowel disease; haematopoietic cell; blood protein disorder;
XX anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
XX endometriosis; pregnancy induced hypertension; ss.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO200061747-A2.
XX
XX PD 19-OCT-2000.
XX
XX 13-APR-2000; 2000WO-US10158.
XX
XX 14-APR-1999; 99US-0129171.
XX 14-MAY-1999; 99US-0134114.
XX 14-MAY-1999; 99US-0134117.
XX 14-MAY-1999; 99US-0134118.
XX 21-OCT-1999; 99US-0160860.
XX 29-OCT-1999; 99US-0162498.
XX 13-DEC-1999; 99US-0170453.
XX 14-JAN-2000; 2000US-0176195.
XX 14-FEB-2000; 2000US-0182296.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX Lu PS:
XX WPI: 2000-619230/59.
XX
XX Isolated cadherin-like asymmetry protein-2 polynucleotide and
XX polypeptide used to diagnose, treat and prevent autoimmune diseases and
XX inflammatory responses -
XX
XX Disclosure: Fig 10F; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antihyaluronic, cytostatic, hypotensive, antirheumatic,
XX antianemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
```


CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g., arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.

XX Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

XX Query Match 100.0%; Score 501; DB 21; Length 4898;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-141;

XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTACACTGGAAGAAGTCTTGTGCGGACACATTTGCAGTCAATCATCTGACGCCA 60
|||
|||

QY 61 GCTGATAGCAGACGTTTGTGCGGAAACAGATTCAGACAGTCCCTGCATCAT 120
|||
|||

Db 2651 GCTGATAGCAGACGTTTGTGCGGAAACAGATTCAGACAGTCCCTGCATCAT 2710
|||
|||

QY 121 CACAACTGTGCGCAACAGTACCGGCTTATTAAAGCACACAGCTTCTCTGATGAA 180
|||
|||

Db 2711 CAAACACTGTGCGCAACAGTACCGGCTTATTAAAGCACACAGCTTCTCTGATGAA 2770
|||
|||

QY 181 GGACTTAACCAAAAGATATGACAGGATGATGACCGCCAGATGAAAGACATGA 240
|||
|||

Db 2771 GGACTTAACCAAAAGATATGACAGGATGATGACCGCCAGATGAAAGACATGA 2830
|||
|||

QY 241 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGGCTGCGCAAAATCCTATGCCAGCAC 300
|||
|||

Db 2831 GAACGACCCAGAGATGCTGTGACCTCCAGTACAGGCTGCGCAAAATCCTATGCCAGCAC 2890
|||
|||

QY 301 GCCCGACTCAGGAAGACGTCGTCGACACATGCGCCAGATCCATGTCAAAAATGGCGA 360
|||
|||

Db 2891 GCCCGACTCAGGAAGACGTCGTCGACACATGCGCCAGATCCATGTCAAAAATGGCGA 2950
|||
|||

QY 361 TCTCTCAGAGGACGATGCTATGTCACGTAACAGCCCTAGTGGCAAAATATCTCAC 420
|||
|||

Db 2951 TCTCTCAGAGGACGATGCTATGTCACGTAACAGCCCTAGTGGCAAAATATCTCAC 3010
|||
|||

QY 421 ACGGAAGGCGTGTATTAGACAGAGATGACCGCTTCAGGGTCATTACCCCAAAATCGA 480
|||
|||

Db 3011 ACGGAAGGCGTGTATTAGACAGAGATGACCGCTTCAGGGTCATTACCCCAAAATCGA 3070
|||
|||

QY 481 CGAGAGGCGCTCCATGATGA 501
|||
|||

Db 3071 CGAGAGGCGCTCCATGATGA 3091
|||
|||

RESULT 15
AAC87980
ID AAC87980 standard; cDNA; 4898 BP.

AC AAC87980;

XX 07-MAR-2001 (first entry)

XX Preliminary CLASP-2 nucleotide sequence #7.

KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;

KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.

XX Homo sapiens.

PN WO200061747-A2.

PD 19-OCT-2000.

PF 13-APR-2000; 2000WO-US10158.

PR 14-APR-1999; 99US-0129171.

PR 14-MAY-1999; 99US-0134114.

PR 14-MAY-1999; 99US-0134117.

PR 14-MAY-1999; 99US-0134118.

PR 21-OCT-1999; 99US-0160860.

PR 29-OCT-1999; 99US-0162498.

PR 13-DEC-1999; 99US-0170453.

PR 14-JAN-2000; 2000US-0176195.

PR 14-FEB-2000; 2000US-0182296.

PA (ARBO-) ARBOR VITA CORP.

XX Lu PS;

XX WPI; 2000-619230/59.

XX Isolated cadherin-like asymmetry protein-2 polynucleotide and
XX polypeptide used to diagnose, treat and prevent autoimmune diseases and
XX inflammatory responses -

XX Disclosure: Fig 10G; 286pp; English.

XX The present invention describes cadherin-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX immune response in a subject may also be inhibited by administering an
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX proteins and antibodies can be used to prevent or treat a CLASP-2
XX mediated disease, such as an autoimmune disease caused or exacerbated
XX by increased activity of TH1 cells. They can also be used to treat
XX hypersensitivities, prevent transplantation rejection responses and
XX augment immune responsiveness in immunodeficiency states, inhibit
XX proliferation and differentiation of cells involved in an inflammatory
XX response e.g., arthritis, inflammatory bowel disease and increase
XX differentiation and proliferation of haematopoietic cells e.g. to treat
XX anaemia, thrombocytopaenia and other blood protein disorders. Disorders
XX treated by disrupting CLASP-2 function include multiple sclerosis,
XX rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
XX The present sequence represents a preliminary CLASP-2 nucleotide
XX sequence, from the present invention.

XX Sequence 4898 BP; 1379 A; 1134 C; 1166 G; 1219 T; 0 other;

XX Query Match 100.0%; Score 501; DB 21; Length 4898;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-141;

XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACACTGGAAGAAGTCTTGTGCGGACACATTTGCAGTCAATCATCTGACGCCA 60
|||
|||

Db 2591 TTACACTGGAAGAAGTCTTGTGCGGACACATTTGCAGTCAATCATCTGACGCCA 2650
|||
|||

QY 61 GCTGATAGCAGACGTTTGTGCGGAAACAGATTCAGACAGTCCCTGCATCAT 120
|||
|||

Db 2651 GCTGATAGCAGACGTTTGTGCGGAAACAGATTCAGACAGTCCCTGCATCAT 2710
|||
|||

QY 121 CACAACTGTGCGCAACAGTACCGGCTTATTAAAGCACACAGCTTCTCTGATGAA 180
|||
|||

```
|||||
Db 2711 CAACACTGTGCCACAGTACGGCTTATTAAGCACACACAGCTTCTCCTGATGTGA 2770
QY 181 GGACTTAACCAAAAGGATACGCACGGTGTATATGGCCACCGCCGAGATGAAGAGCATGA 240
|||||
Db 2771 GGACTTAACCAAAAGGATACGCACGGTGTATATGGCCACCGCCGAGATGAAGAGCATGA 2830
QY 241 GAACGACCCAGAGATGCTGTGGTACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
|||||
Db 2831 GAACGACCCAGAGATGCTGTGGTACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890
QY 301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGATCCATGTCAAAAATGGCGA 360
|||||
Db 2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGATCCATGTCAAAAATGGCGA 2950
QY 361 TCTCTCAGAGGACGATGTGCTATGTCACGTAACAGCCCTAGTGGCAGATATCTCAC 420
|||||
Db 2951 TCTCTCAGAGGACGATGTGCTATGTCACGTAACAGCCCTAGTGGCAGATATCTCAC 3010
QY 421 ACGGAAGGCGTGTTAGACAGATGCACGCTTCAGAGGTCAATACCCCAACATCGA 480
|||||
Db 3011 ACGGAAGGCGTGTTAGACAGATGCACGCTTCAGAGGTCAATACCCCAACATCGA 3070
QY 481 CGAGGAGGCGCTCCATGATGGA 501
|||||
Db 3071 CGAGGAGGCGCTCCATGATGGA 3091
```

Search completed: February 7, 2003, 07:08:07
Job time : 158.085 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:00:34 ; Search time 28.6858 Seconds
(Without alignments)
5356.145 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000

Perfect score: 501
Sequence: 1 ttactctggaagaagtcct.....gaggaagcctcatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	8.6	4722	4	US-08-979-608A-14
2	37.4	7.5	1432	1	US-07-914-282D-3
3	37.4	7.5	1432	1	US-08-276-887A-3
4	37.4	7.5	1432	5	PCT-US93-02460-3
5	36.4	7.3	1111	2	US-08-997-080-169
6	36.4	7.3	1111	2	US-08-997-362-169
7	36.4	7.3	1111	4	US-09-095-855-169
8	36.4	7.3	1111	4	US-09-324-542-169
9	36.4	7.3	1111	4	US-09-205-426-169
10	36.4	7.3	1341	2	US-08-997-080-93
11	36.4	7.3	1341	2	US-08-997-362-93
12	36.4	7.3	1341	3	US-08-873-970-93
13	36.4	7.3	1341	4	US-09-095-855-93
14	36.4	7.3	1341	4	US-09-324-542-93
15	36.4	7.3	1341	4	US-09-205-426-93
16	35	7.0	1854	5	PCT-US94-01101-1
17	35	7.0	4697	4	US-08-979-608A-17
18	35	6.7	28171	4	US-08-961-527-22
19	32.6	6.5	4972	3	US-09-035-648-17
20	32.6	6.5	4972	4	US-09-001-951-17
21	32.6	6.5	4972	4	US-08-818-829-17
22	32.2	6.4	1136	1	US-08-106-981-5
23	32.2	6.4	4403765	4	US-09-103-840A-2
24	32.2	6.4	4411529	4	US-09-103-840A-1
25	32	6.4	2419	4	US-09-245-041-8
26	32	6.4	8827	4	US-09-245-041-1
27	32	6.4	62804	4	US-09-800-960-3

28	31.8	6.3	4443	4	US-09-425-453A-1	Sequence 1, Appl
29	31.8	6.3	4443	4	US-09-425-453A-3	Sequence 3, Appl
30	31.8	6.3	4443	4	US-09-425-453A-5	Sequence 5, Appl
31	31.8	6.3	4443	4	US-09-425-453A-7	Sequence 7, Appl
32	31.8	6.3	4443	4	US-09-425-453A-9	Sequence 9, Appl
33	31.8	6.3	4443	4	US-09-425-453A-11	Sequence 11, Appl
34	31.8	6.3	4443	4	US-09-425-453A-13	Sequence 13, Appl
35	31.8	6.3	4443	4	US-09-425-453A-15	Sequence 15, Appl
36	31.8	6.3	4443	4	US-09-425-453A-17	Sequence 17, Appl
37	31.8	6.3	4443	4	US-09-425-453A-19	Sequence 19, Appl
38	31.8	6.3	4560	4	US-09-256-703-1	Sequence 1, Appl
39	31.8	6.3	5635	1	US-08-136-742A-3	Sequence 3, Appl
40	31.8	6.3	5635	3	US-09-248-026-3	Sequence 3, Appl
41	31.8	6.3	5635	5	PCT-US93-11667-3	Sequence 3, Appl
42	31.8	6.3	6126	2	US-08-951-912-3	Sequence 3, Appl
43	31.8	6.3	6126	4	US-09-174-077-3	Sequence 3, Appl
44	31.8	6.3	6129	1	US-07-637-621-1	Sequence 1, Appl
45	31.8	6.3	6129	1	US-08-136-742A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-979-608A-14
; Sequence 14, Application US/08979608A
; Patent NO. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Law, Robert S.
; Atjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEIN AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979, 608A
; FILING DATE: 26-NOV. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048, 547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031, 930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35, 965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...1731
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 1111 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-095-855-169

Query Match
Best Local Similarity 46.2%; DB 4; Length 1111;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAGCAGTCCCTGTCATATTAACAACACTGTGCCAAGAGTACCGGCTTTATAGCACAC 159
DB 805 CGAGGGGTGATGACTACTATAGACCGAGCCAGCACTAGCCCAAGTGTCCGTTCCAC 864
QY 160 CACCTTCTCTCTGATGTGAAGACTTAACCAAAAGATACGACGAGTGTGCTAATGCGCAC 219
DB 865 CCAAGTTCGTGCGCCGACCTCGCATGACCGAGAACTGCCCAAGTGCATCTGCATC 924
QY 220 CGCCCGATGAAGAGCATGAGACGACCCAGAGATCTGTGTGACCTCCAGTACAGCCT 279
DB 925 GGGGAGAACCCGCTGATCAACCCCTCGCGAGGTGAGGCGAAGCTGAAGTGTGGGC 984
QY 280 GGCCAATCTATGCGCAGCAGCCCGAGCTGAGAAAGCTGGCTGCACAGCATGGCCAG 339
DB 985 GGCACGTGACCGAGCAGACAGCCAGAGATTCAACACTGCTACGCCGCGCTACCGCGG 1044
QY 340 GATCCATGTCAAAAATGGCGAT 361
DB 1045 CTGACGCGGTGTAGTCCGAT 1066

RESULT 8

US-09-324-542-169
Sequence 169, Application US/09324542
Patent No. 6328978
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L.J.
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Methods and Compounds for the Treatment
TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
FILE REFERENCE: 11000.1007c1
CURRENT APPLICATION NUMBER: US/09/324,542
CURRENT FILING DATE: 1999-06-02
EARLIER APPLICATION NUMBER: US 08/997,080
EARLIER FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 194
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 169
LENGTH: 1111
TYPE: DNA
ORGANISM: Mycobacterium vaccae
US-09-324-542-169

Query Match
Best Local Similarity 7.3%; Score 36.4; DB 4; Length 1111;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAGCAGTCCCTGTCATATTAACAACACTGTGCCAAGAGTACCGGCTTTATAGCACAC 159
DB 805 CGAGGGGTGATGACTACTATAGACCGAGCCAGCACTAGCCCAAGTGTCCGTTCCAC 864
QY 160 CACCTTCTCTCTGATGTGAAGACTTAACCAAAAGATACGACGAGTGTGCTAATGCGCAC 219
DB 865 CCAAGTTCGTGCGCCGACCTCGCATGACCGAGAACTGCCCAAGTGCATCTGCATC 924
QY 220 CGCCCGATGAAGAGCATGAGACGACCCAGAGATCTGTGTGACCTCCAGTACAGCCT 279
DB 925 GGGGAGAACCCGCTGATCAACCCCTCGCGAGGTGAGGCGAAGCTGAAGTGTGGGC 984
QY 280 GGCCAATCTATGCGCAGCAGCCCGAGCTGAGAAAGCTGGCTGCACAGCATGGCCAG 339

DB 985 GGCACGTGACCGAGCAGACAGAGATTCAACACTGCTACGCCGCGCTACCGCGG 1044
QY 340 GATCCATGTCAAAAATGGCGAT 361
DB 1045 CTGACGCGGTGTAGTCCGAT 1066

RESULT 9

US-09-205-426-169
Sequence 169, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L.J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
FILE REFERENCE: 11000.1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 169
LENGTH: 1111
TYPE: DNA
ORGANISM: Mycobacterium vaccae
US-09-205-426-169

Query Match
Best Local Similarity 7.3%; Score 36.4; DB 4; Length 1111;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAGCAGTCCCTGTCATATTAACAACACTGTGCCAAGAGTACCGGCTTTATAGCACAC 159
DB 805 CGAGGGGTGATGACTACTATAGACCGAGCCAGCACTAGCCCAAGTGTCCGTTCCAC 864
QY 160 CACCTTCTCTCTGATGTGAAGACTTAACCAAAAGATACGACGAGTGTGCTAATGCGCAC 219
DB 865 CCAAGTTCGTGCGCCGACCTCGCATGACCGAGAACTGCCCAAGTGCATCTGCATC 924
QY 220 CGCCCGATGAAGAGCATGAGACGACCCAGAGATCTGTGTGACCTCCAGTACAGCCT 279
DB 925 GGGGAGAACCCGCTGATCAACCCCTCGCGAGGTGAGGCGAAGCTGAAGTGTGGGC 984
QY 280 GGCCAATCTATGCGCAGCAGCCCGAGCTGAGAAAGCTGGCTGCACAGCATGGCCAG 339
DB 985 GGCACGTGACCGAGCAGACAGCGAGATTCAACACTGCTACGCCGCGCTACCGCGG 1044
QY 340 GATCCATGTCAAAAATGGCGAT 361
DB 1045 CTGACGCGGTGTAGTCCGAT 1066

RESULT 10

US-08-997-080-93
Sequence 93, Application US/08997080
Patent No. 5968524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 1341 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-997-080-93

Query Match 7.3%; Score 36.4; DB 2; Length 1341;
Best Local Similarity 46.2%; Pred. No. 0.072;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 GCACGATCCCTGTCATCAACAACTGTGCCAAGTGGCGCTTATTAAACAC 159
DB 979 CGAGGCGTGATCGACTCATCTACGACCGAAGCAAGCTGGTGGCTTAC 1038
QY 160 CAGCTTCTCTGTATGTGAAGACTTAACCAAAAGATACGACGGTCTAATGGCCAC 219
DB 1039 CCAATTCGTGCGCCGACCTCGACATGACGACCACTCCCAAGGTGATCTCTGCATC 1098
QY 220 GCGCCAGATGAAGAGCATGAAGACGACCGAGATGCTGGAGACCTCCAGTACAGCCT 279
DB 1099 GCGGAGAACCCGCTGATCAACCCGTGCGCGAGGTGACAGGGAACCTGAAAGTGTGGGC 1158
QY 280 GCGCAATTCATGCCAGACGCGCCAGCTCAGAAAGATGCTGACAGATGGCCAG 339
DB 1159 GCGACTGACCGACGAGACGACGAGAGTTCAACACTGGTACGCCGCGTACACGGCGG 1218
QY 340 GATCCATGTCAAAATGGCGAT 361
DB 1219 CTGACGCGGTGTAGTCCGAT 1240

RESULT 11
US-08-997-362-93
Sequence 93, Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Ann W. Speckman
SHEET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 1341 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-997-362-93

Query Match 7.3%; Score 36.4; DB 2; Length 1341;
Best Local Similarity 46.2%; Pred. No. 0.072;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAAGATCCCTGTCATCAACAACTGTGCCAAGTGGCGCTTATTAAACAC 159
DB 979 CGAGGCGTGATGATCACTCTACGACCGAAGCAAGCTGGTGGCTTAC 1038
QY 160 CAGCTTCTCTGTATGTGAAGACTTAACCAAAAGATACGACGGTCTAATGGCCAC 219
DB 1039 CCAATTCGTGCGCCGACCTCGACATGACGACCACTCCCAAGGTGATCTCTGCATC 1098
QY 220 GCGCCAGATGAAGAGCATGAGAAGACGACCGAGATGCTGGAGCTCCAGTACAGCCT 279
DB 1099 GCGGAGAACCCGCTGATCAACCCGTGCGCGAGGTGACAGGGAACCTGAAAGTGTGGGC 1158
QY 280 GCGCAATTCATGCCAGACGCGCCAGCTCAGAAAGATGCTGACAGATGGCCAG 339
DB 1159 GCGACTGACCGACGAGACGACGAGAGTTCAACACTGGTACGCCGCGTACACGGCGG 1218
QY 340 GATCCATGTCAAAATGGCGAT 361
DB 1219 CTGACGCGGTGTAGTCCGAT 1240

RESULT 12
US-08-873-970-93
Sequence 93, Application US/08873970
Patent No. 6001361
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross


```

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-873-970-93

Query Match
Best Local Similarity 7.3%; Score 36.4; DB 3; Length 1341;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAGCAGTCCCTGTCATCATCAACAACGTGCGCAACAGTACCGGCTTATTAAACACAC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 979 CGAGGCGTGATGATCATCTACATCTACGACGAGCGCAACTAGCCAGGTGTCGGCTTAC 1038

QY 160 CAGTTCTCTCTATGTGAAGACTTAACCAAAAGGATACGACGGTCTAATGGCCAC 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1039 CCACTTGTGCCCGCACTCTCGACATGACCGAGCAACTCGCCAGGTGCTGATCTGCATC 1098

QY 220 CGCCGATGAAGAGGATGAGAACGACCGACGATGCTGGTGAGCTTCAGTACAGCT 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1099 GCGGGAACACCCGCTGATCAACCCGCTGGCGAGGTGACGCAACCTGAAGTCGTGGGC 1158

QY 280 GGCCTTCTATGTCAGACGCGCCGAGCTCAGGAAGACGTGCTGACACATGGCCAG 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1159 GGCACGTACCGAGCAGACGACGAGGAGTTCAACACTCGTACGCGCGCTCACCGGCGG 1218

QY 340 GATCATGTCAAAAATGGCGAT 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1219 CTGACGCGGTGTAGTCCGAT 1240

RESULT 13
US-09-095-855-93
; Sequence 93, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
```

```

; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-09-095-855-93

Query Match
Best Local Similarity 7.3%; Score 36.4; DB 4; Length 1341;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAGCAGTCCCTGTCATCATCAACAACGTGCGCAACAGTACCGGCTTATTAAACACAC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 979 CGAGGCGTGATGATCATCTACATCTACGACGAGCGCAACTAGCCAGGTGTCGGCTTAC 1038

QY 160 CAGTTCTCTCTATGTGAAGACTTAACCAAAAGGATACGACGGTCTAATGGCCAC 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1039 CCACTTGTGCCCGCACTCTCGACATGACCGACGCAACTCGCCAGGTGCTGATCTGCATC 1098

QY 220 CGCCGATGAAGAGGATGAGAACGACCGACGATGCTGGTGAGCTTCAGTACAGCT 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1099 GCGGGAACACCCGCTGATCAACCCGCTGGCGAGGTGACGCAACCTGAAGTCGTGGGC 1158

QY 280 GGCCTTCTATGTCAGACGCGCCGAGCTCAGGAAGACGTGCTGACACATGGCCAG 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1159 GGCACGTACCGAGCAGACGACGAGGAGTTCAACACTCGTACGCGCGCTCACCGGCGG 1218

QY 340 GATCATGTCAAAAATGGCGAT 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1219 CTGACGCGGTGTAGTCCGAT 1240

RESULT 14
US-09-324-542-93
; Sequence 93, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
```

APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L.J.
TITLE OF INVENTION: Methods and Compounds for the Treatment
TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
FILE REFERENCE: 11000.1007c1
CURRENT APPLICATION NUMBER: US/09/324,542
EARLIER APPLICATION NUMBER: US 08/997,080
EARLIER FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 194
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 93
LENGTH: 1341
TYPE: DNA
ORGANISM: Mycobacterium vaccae
US-09-324-542-93

Query Match 7.3%; Score 36.4; DB 4; Length 1341;
Best Local Similarity 46.2%; Pred. No. 0.072;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 100 CCAGCAGTCCCTGTCCATCATCAACACTGTGCCAACAGTGCAGCGCTTATTAAAGCACAC 159
DB 979 CGAGCGCTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGTGCGCTTCAC 1038
QY 160 CAGCTTCTCCTGTGATGTGAGGACTTAAACCAAAAGGATACGACGCTTAATGGCCAC 219
DB 1039 CCAGTTCTGCTCCGCTGCTGACATGACCGACGACGACGACGACGACGACGACGACGAC 1098
QY 220 CGCCAGATGAAGAGCATGAGAACGCCAGAGATGCTGTGACCTCCAGTACAGCCT 279
DB 1099 GCGGAGAACCCCGCTGATCAACCCCGTGGCGGAGTGCAAGGCAACCTGAAGTGTGGGC 1158
QY 280 GGCCTAATCTATGCGACAGCGCCGAGCTCAGGAAAGCTGCTGACAGCATGGCCAG 339
DB 1159 GGCACCTGACGACGAGAGAGGAGTTCACACACTGCTGACGCGCGCTACCGCGCG 1218
QY 340 GATCCATGTCAAAAATGGCGAT 361
DB 1219 CTGACGCGGTGTGTGCTCGAT 1240

RESULT 15
US-09-205-426-93
Sequence 93, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
FILE REFERENCE: 11000.1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 93
LENGTH: 1341
TYPE: DNA
ORGANISM: Mycobacterium vaccae
US-09-205-426-93

Query Match 7.3%; Score 36.4; DB 4; Length 1341;
Best Local Similarity 46.2%; Pred. No. 0.072;

Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
QY 100 CCAGCAGTCCCTGTCCATCATCAACACTGTGCCAACAGTGCAGCGCTTATTAAAGCACAC 159
DB 979 CGAGCGCTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGTGCGCTTCAC 1038
QY 160 CAGCTTCTCCTGTGATGTGAGGACTTAAACCAAAAGGATACGACGCTTAATGGCCAC 219
DB 1039 CCAGTTCTGCTCCGCTGCTGACATGACCGACGACGACGACGACGACGACGACGACGAC 1098
QY 220 CGCCAGATGAAGAGCATGAGAACGCCAGAGATGCTGTGACCTCCAGTACAGCCT 279
DB 1099 GCGGAGAACCCCGCTGATCAACCCCGTGGCGGAGTGCAAGGCAACCTGAAGTGTGGGC 1158
QY 280 GGCCTAATCTATGCGACAGCGCCGAGCTCAGGAAAGCTGCTGACAGCATGGCCAG 339
DB 1159 GGCACCTGACGACGAGAGGAGTTCACACACTGCTGACGCGCGCTACCGCGCG 1218
QY 340 GATCCATGTCAAAAATGGCGAT 361
DB 1219 CTGACGCGGTGTGTGCTCGAT 1240

Search completed: February 7, 2003, 09:05:47
Job time : 43.6858 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:01:14 ; Search time 32.3549 Seconds
(without alignments)
7339.716 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000

Perfect score: 501

Sequence: 1 ttacactggaagaagctct.....gagagagcctcatgatgga 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapect 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCF_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCFUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	274.4	54.8	6454	10 US-09-736-969A-1	Sequence 1, Appl1
2	269.6	53.8	4391	10 US-09-736-969A-7	Sequence 7, Appl1
3	130.6	26.1	483	10 US-09-864-761-13960	Sequence 13960, A
4	116.4	23.2	2149	9 US-09-736-968A-59	Sequence 59, Appl1
5	116.4	23.2	6372	9 US-09-736-968A-1	Sequence 1, Appl1
6	115.4	23.0	212	10 US-09-864-761-30524	Sequence 30524, A
7	83.2	16.6	4026	10 US-09-736-960-3	Sequence 3, Appl1
8	83.2	16.6	7215	10 US-09-736-960-1	Sequence 1, Appl1
9	79.4	15.8	427	9 US-09-796-692-5465	Sequence 5465, Ap
10	69	13.8	444	9 US-09-796-692-3710	Sequence 3710, Ap
11	57.4	11.5	271	9 US-09-736-968A-98	Sequence 98, Appl1
12	57.4	11.5	512	10 US-09-864-761-7634	Sequence 7634, Ap
13	50	10.0	194	10 US-09-864-761-24339	Sequence 24339, A
14	48.8	9.7	266	10 US-09-736-960-76	Sequence 76, Appl1
15	48.8	9.7	6686	10 US-09-736-960-86	Sequence 86, Appl1
16	46.2	9.2	379	9 US-10-046-935-2206	Sequence 2206, Ap
17	46.2	9.2	379	9 US-09-878-178-2206	Sequence 2206, Ap
18	45.4	9.1	181	9 US-09-736-968A-97	Sequence 97, Appl1
19	43	8.6	4722	9 US-09-976-740-14	Sequence 14, Appl1

20	43	8.6	4722	10 US-09-962-055-14	Sequence 14, Appl1
21	43	8.6	4722	12 US-10-023-529-14	Sequence 14, Appl1
22	43	8.6	4722	12 US-10-023-523-14	Sequence 14, Appl1
23	36.4	7.3	1111	9 US-10-051-643-169	Sequence 169, Appl
24	36.4	7.3	1111	9 US-09-880-505-169	Sequence 169, Appl
25	36.4	7.3	1341	9 US-10-051-643-93	Sequence 93, Appl1
26	36.4	7.3	1341	9 US-09-880-505-93	Sequence 93, Appl1
27	35.4	7.1	292	10 US-09-815-242-872	Sequence 872, App
28	35.4	7.1	405	10 US-09-815-242-3815	Sequence 3815, Ap
29	35.4	7.1	435	10 US-09-815-242-6463	Sequence 6463, Ap
30	35.4	7.1	558	10 US-09-815-242-112	Sequence 112, App
31	35.4	7.1	558	10 US-09-815-242-133	Sequence 133, App
32	35.4	7.1	583	10 US-09-815-242-147	Sequence 147, App
33	35.4	7.1	14557	10 US-09-070-927A-367	Sequence 367, App
34	35	7.0	1638	9 US-09-976-740-46	Sequence 46, Appl1
35	35	7.0	1638	12 US-10-023-529-46	Sequence 46, Appl1
36	35	7.0	1638	12 US-10-023-523-46	Sequence 46, Appl1
37	35	7.0	4697	9 US-09-976-740-17	Sequence 17, Appl1
38	35	7.0	4697	10 US-09-962-055-17	Sequence 17, Appl1
39	35	7.0	4697	12 US-10-023-529-17	Sequence 17, Appl1
40	35	7.0	4697	12 US-10-023-523-17	Sequence 17, Appl1
41	33.6	6.7	891	10 US-09-815-242-9579	Sequence 9579, Ap
42	32.8	6.5	88421	9 US-09-976-059-1	Sequence 1, Appl1
43	32.2	6.4	2133	9 US-10-124-800-17	Sequence 17, Appl1
44	32.2	6.4	2930	10 US-09-745-763-198	Sequence 198, Appl
45	32.2	6.4	8730	9 US-10-124-800-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

US-09-736-969A-1

Sequence 1, Application US/09736969A

Patent No. US20020068302A1

GENERAL INFORMATION:

APPLICANT: Lu, Peter

APPLICANT: Garman, Jonathan David

APPLICANT: Candia IT, Albert Frederick

APPLICANT: Arbor Vita Corporation

TITLE OF INVENTION: CLASP-4 Transmembrane Protein

FILE REFERENCE: 020054-00041US

CURRENT APPLICATION NUMBER: US/09/736,969A

CURRENT FILING DATE: 2000-12-13

PRIOR APPLICATION NUMBER: US 60/160,860

PRIOR FILING DATE: 1999-10-21

PRIOR APPLICATION NUMBER: US 60/162,498

PRIOR FILING DATE: 1999-10-29

PRIOR APPLICATION NUMBER: US 60/170,453

PRIOR FILING DATE: 1999-12-13

PRIOR APPLICATION NUMBER: US 60/176,195

PRIOR FILING DATE: 2000-01-14

PRIOR APPLICATION NUMBER: US 60/182,296

PRIOR FILING DATE: 2000-02-14

PRIOR APPLICATION NUMBER: US 09/547,276

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: US 60/196,267

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: US 60/196,460

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: US 60/196,527

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: US 60/196,528

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: US 09/687,837

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: US 60/240,503

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: US 60/240,508

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: US 60/240,539

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: US 60/240,543

;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 153
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 1
;; LENGTH: 6454
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: full length human CLASP-4 cDNA
;; NAME/KEY: CDS
;; LOCATION: (95)..(6121)
;; OTHER INFORMATION: human CLASP-4
US-09-736-969A-1

Query Match 54.8%; Score 274.4; DB 10; Length 6454;
Best Local Similarity 71.8%; Pred. No. 2.6e-74;
Matches 359; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 2 TACACTGGAAGAAGTCTTTGTCGGACACATTTCAGATCATCATCTGTGCAGCAG 61
Db 4442 TATACCAAGGAAGAACTTTTGAGACACATCTACAGATATATATGCTGTGAAGCCA 4501
QY 62 CTGATAGCAGACGTTGTGGCATTTGGGGAACCCAGATTCAGAGTCCCTGCATCATC 121
Db 4502 CTGATAGCTGATGAGACTAGACGAGGATCAAGATTTCAGAGTCTTATTCATTATC 4561
QY 122 AACACAGTGTCCCAAGATGACCGGCTTATTAAGCACACCAGCTTCCTCTGATGTGAAG 181
Db 4562 AATAAATTTTGCATAATGTGACAGACCTATGAAAGCAATGCCCTTCCCGAGAAGTCAAA 4621
QY 182 GACTTAACCAAGAAGATACGACGCTGTATATGAGCCACCGCCAGATGAAGAGCATGAG 241
Db 4622 GACTTGACCAAGAAATCCGCACTGTTCTTATGCCCCATGCCCCAATGAAGAGCATGAG 4681
QY 242 AACGACCCAGAGATGCTGTGACCTCCAGTACAGCCTGCGCAATCTATGCGACAG 301
Db 4682 AAGACCCCTGAATGCTAATTTGATGTCCAGTATAGCTTAGCCAAAGCTATGCAAGCAC 4741
QY 302 CCGGAGCTCGAAGAAGATGCTGCTGACAGCTGACAGATGCGCAGATCATGTCAAAATGGCGAT 361
Db 4742 CCGAGGCTCGAAGAAGCTGCTGATGATGCGCAAGATTCATTAATAAATGGAGAT 4801
QY 362 CTCTCAGAGCAGCAATGTCTATGTCCAGTACAGACCTTAGTAGCAGATATCTCACA 421
Db 4802 TTTTCAGAGGCTCGATGTGTATGTCCAGTACAGCTCTAAGATTTCTTCAT 4861
QY 422 CGGAAGGCGTGTTTAGACAGAGTGCACCGCCTTCAGGCTCATTAACCCAAACATGCAC 481
Db 4862 CGAAGAAATTTATTTCTTAAGGATGTTCAGCCTTCAGAAATTAATCTCCCAATATGAT 4921
QY 482 GAGGAGGCTCCATGATGCA 501
Db 4922 GAAGAAGGAGCAATGAAGA 4941

RESULT 2

US-09-736-969A-7
; Sequence 7, Application US/09736969A
; Patent No. US20020068302N1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia, III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736,969A
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29

;; PRIOR APPLICATION NUMBER: US 60/170,453
;; PRIOR FILING DATE: 1999-12-13
;; PRIOR APPLICATION NUMBER: US 60/176,195
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 60/182,296
;; PRIOR FILING DATE: 2000-02-14
;; PRIOR APPLICATION NUMBER: US 09/547,276
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,267
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,460
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,527
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 60/196,528
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: US 09/687,837
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,503
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,508
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,539
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/240,543
;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 153
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 7
;; LENGTH: 4391
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: human CLASP-4 cDNA
;; NAME/KEY: CDS
;; LOCATION: (414)..(4058)
;; OTHER INFORMATION: human CLASP-4
US-09-736-969A-7

Query Match 53.8%; Score 269.6; DB 10; Length 4391;
Best Local Similarity 71.2%; Pred. No. 6.6e-73;
Matches 356; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 2 TACACTGGAAGAAGTCTTTGTCGGACACATTTCAGATCATCATCTGTGCAGCAG 61
Db 2379 TATACCAAGGAAGAACTTTTGAGACACATCTACAGATATATATGCTGTGAAGCCA 2438
QY 62 CTGATAGCAGACGTTGTGGCATTTGGGGAACCCAGATTCAGAGTCCCTGCATCATC 121
Db 2439 CTGATAGCTGATGTAGCACTTAAGCGGAGAGATCAAGATTTCAGAGTCTTATTCATTATC 2498
QY 122 AACACAGTGTCCCAAGATGACCGGCTTATTAAGCACACCAGCTTCCTCTGATGTGAAG 181
Db 2499 AATAAATTTTGCATAATGTGACAGACCTATGTGGCAAGAGCCTTCCCGAGAAGTCAAA 2558
QY 182 GACTTAACCAAGAAGATACGACGCTGTATATGAGCCACCGCCAGATGAAGAGCATGAG 241
Db 2559 GACTTGACCAAGAAATCCGCACTGTTCTTATGTAGGCACTGCCCCAAAGAGAGAGCATGAG 2618
QY 242 AACGACCCAGAGATGCTGTGACCTCCAGTACAGCCTTGCCAAATCTATGCGACAGC 301
Db 2619 AAGAGCCCTGAATGTCTATTTGATCTCCAGTATAGCTTAGCAGATGCTTATGCAAGCACC 2678
QY 302 CCGGAGCTCGAAGAAGCTGCTGACAGCATGAGCCAGAGATTCATGTCAAAATGGCGAT 361
Db 2679 CCAGAGCTCGAAGAAACCTGCTGATGATGATGAGCATGAGCAAGATTCATGAAGAGAGAT 2738
QY 362 CTCTCAGAGGAGCAAGTGTCTATGTCCAGTAAACAGCCCTAGTGGAGAAATATCTACA 421
Db 2739 TTTTCAGAGGCTCGATGTGTATGTCCATGTACAGCTCTCATTTGAGAGTGTCTTCAT 2798
QY 422 CGGAAGGCGTGTTTAGACAGAGATGACCGCCTTCAGAGGCTCATTAACCCAAACATGCAC 481

Db 2799 CGAAGAAAATATTTCTTAACGAGATGTTCAAGCTTCAAGAAAATTAATCTCCCAATATAGAT 2858
Qy 482 GAGAGAGCCTCCATCATGATGA 501
Db 2859 GAAGAGAGACATGAAGA 2878

RESULT 3

US-09-864-761-13960
; Sequence 13960, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13960
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011739.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
US-09-864-761-13960

Query Match 26.1%; Score 130.6; DB 10; Length 483;
Best Local Similarity 78.1%; Pred. No. 1.2e-30;
Matches 157; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 171 CTGATGTGAAGAGCTTAACCAAAAGATACGACGCTCTAATGCCCCACATGA 230
Db 269 CAGAGTGAAGAGACCTGATTAAGCCGTATTAAGCATGTTTATGATGGCCACAGCTCAGATGA 328
Qy 231 AGGAGCATGAGAACGACGACGAGATGCTGTGAGCTCCAGTACAGCTGACCAAAATCCT 290
Db 329 AGGAGCAGAGAGAGACCCGAGATGCTGTGATCTCAATACAGCTG6CAAACTCCT 388
Qy 291 ATGCCACAGCCCCGAGCTCAGAGACGTGGCTGACAGCATGGCCAGATCCATGTCA 350
Db 389 AGCGAACACCTCCTGAACTACGACGACCTGGCTGAAAGTATG6CCAAATTCATGCCA 448
Qy 351 AAAATGGAGATCTCTCAGAGG 371
Db 449 GAACGGAGATTTATCTGAGG 469

RESULT 4

US-09-736-968A-59
; Sequence 59, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Gorman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbot Vita Corporation
; TITLE OF INVENTION: CLASP-7 Transmembrane Protein
; FILE REFERENCE: 020054-00061105
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US/09/736,968A
; PRIOR FILING DATE: 2000-10-21
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 2149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: preliminary human CLASP-7 cDNA sequence
; NAME/KEY: CDS
; LOCATION: (2)..(1933)
; OTHER INFORMATION: human CLASP-7
US-09-736-968A-59

```

PRIORITY FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,543
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 115
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 6372
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: full length human CLASP-7 cDNA
NAME/KEY: CDS
LOCATION: (13)..(6156)
OTHER INFORMATION: human CLASP-7
US-09-736-968A-1

Query Match          23.2%   Score 116.4; DB 9: Length 6372;
Best Local Similarity 55.4%; Pred. No. 9, 2e-26;
Matches 225; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

OY      19    CTTTTCGCGACATTTGCAGTAGCATATCTGTCAAGCCACTGATAGCAGACTGGT 78
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4506  CTATTGCCCTGTGAAMATCACAGTACCAGTCTCTCTCGCTCCCTGGTGCGAGCAGCA 4565

OY       79    TGGCATTTGGGGAAGAACAGATTCACAGAGCCCTGTCATCAACAGTGTCCAACAG 138
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     4566  GAACATTAGAGAAGACACCTCTGCAAGTTCACATAAAACATCTCTACCTGACTAGGA 4625

OY      139    TGACCCGCTTAATTAAACAACAACAGCTTCTCCTCTGATGTGAAGACTTTAAACAAAAGAT 198
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     4626  GGCAATCGGGGCTCCGGGACAGACCTTGCGACAGCAGGTCCAGGACCTGATGTTCAACT 4685

OY      199    ACGCAGCGTGCTATATGAGCCAGCCAGATGAAGAGAGCATGANAAGCCAGATGATGCT 258
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     4686  GCACATGATCCTGACGAGACAGGTGAAGTGAAGGAACACAGGAGGACCTGATGATGCT 4745

OY      259    GGTTGACCTCAATACAGCCTTGCCCAATCTCTATGCCAGACAGCGCCGAGCTCAGGAAGAC 318
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     4746  CATGCACTTATATACAGATTGCCCCGGGGCTTACCAAGGGCTCACCAGGACCTTGAGCTGAC 4805

OY      319    GTGGCTTCGACAGATGCCCAGAGATCCATGTCAAATAATGGCGATCTTCAAGAGGACCAAT 378
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     4806  CTGGTTCAGAACATGATGGCCGGGAACACGCGAGAGCTGGCAACCAACGACGCGCCCA 4865

OY      379    GTGCTATGTCAGCTAACAGCCCTAGTGGCAAGATATCTCAACAGG 424
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     4866  GTGCATGTGTCACGCGGCCCTCTGCTGCTGATGACTGACCTG 4911

RESULT 6
US-09-864-761-30524
Sequence 30524, Application US/09864761
Patent NO. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263..6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
```

PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30524
LENGTH: 212
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011739.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
OTHER INFORMATION: EST_HUMAN HIT: BE896788.1, EVALUE 9.00e-93
OTHER INFORMATION: SWISSPROT HIT: O24696, EVALUE 3.90e-01
OTHER INFORMATION: NT HIT: g11429300, EVALUE 6.00e-93
US-09-864-761-30524

Query Match 23.0%; Score 115.4; DB 10; Length 212;
Best Local Similarity 79.2%; Pred. No. 4e-26; Indels 0; Gaps 0;
Matches 137; Conservative 0; Mismatches 36;

QY 199 ACGCAGCGTCTAATGAGCCACCCGAGATGAAGAGCATGGAAGAGAGATGCT 258
DB 1 AAGGACTGTTTGTGATGGCCACAGCTCAGATGAAGAGAGAGAGAGAGATGCT 60
QY 259 GGTGAGACTCCAGTACAGCTGGCCAAATCTATGCGACAGCCCGAGCTAGAGAAGC 318
DB 61 GGTGATCTCAGTACAGCTGGCCAAATCTATGCGACAGCAGCTCTGAACTACGAGAGC 120
QY 319 GTGCGCTGACAGCATGGCCAGATCCATGTCAAAATGCGCATCTCTCAGAG 371
DB 121 CTGGCTGTGAAGATATGGCCAGATTCATGCGCAAGAGATTTATCTGAGG 173

RESULT 7
US-09-736-960-3
Sequence 3, Application US/09736960
Patent No. US20020102267A1
GENERAL INFORMATION:
APPLICANT: Lu, Peter
APPLICANT: Garmen, Jonathan David
APPLICANT: Candia III, Albert Frederick
TITLE OF INVENTION: Arbor Vita Corporation
FILE REFERENCE: 020054-000511US
CURRENT APPLICATION NUMBER: US/09/736,960
PRIOR APPLICATION NUMBER: US 60/160,860
PRIOR FILING DATE: 1999-10-21
PRIOR APPLICATION NUMBER: US 60/162,498

PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: US 60/170,453
PRIOR FILING DATE: 1999-12-13
PRIOR APPLICATION NUMBER: US 60/176,195
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: US 60/182,296
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: US 09/547,276
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,267
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,460
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,527
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,528
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 09/687,837
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,503
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,508
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,539
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,543
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 134
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 4026
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: preliminary human CLASP-5 cDNA sequence
NAME/KEY: CCS
LOCATION: (1)..(3066)
OTHER INFORMATION: human CLASP-5
US-09-736-960-3

Query Match 16.6%; Score 83.2; DB 10; Length 4026;
Best Local Similarity 50.5%; Pred. No. 1.2e-15;
Matches 202; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 20 TTGTCCGACACATTTGACATCATATCTGCAGCCAGCTGATAGCAGAGCTGTT 79
DB 1300 TTGGCAAGAGTAAGATGCAAGTAACCATGCTCGGAGATCTTGTGGGAAGACACCA 1359
QY 80 GGCATTGGGAAACCGAGATTCCAGAGTCCCTGTCCATCATCAACAACTGTGCCAAGT 139
DB 1360 GACTTTAATGAAGAGCACTGAGAGAGATCTTGTAGACAAATTTGGCTATTCAGAAAG 1419
QY 140 GACCGGCTTATTAAGACACAGCTTCTCTGATGTGAAGAGCTTAACAAAGATA 199
DB 1420 GACACAGCCATGAGATGACTCTTTCCACAGGTGAGAGACTTCTGTATATGTG 1479
QY 200 CGCAGGTGCTAATGCGCCACCGCCAGATGAAGAGAGATGAGACAGACCCAGAGTCTG 259
DB 1480 AATAGCATCTTATATGACACAGTGAAGAAATGAGGAAATTTCAAGAAAGATCTCAGATGCTT 1539
QY 260 GTGAGCTTCAGATGACAGCTGGCCAAATCTTATGCGACAGCCCGAGCTGGAAGAGC 319
DB 1540 ATGATCTCATGTACAGATGAGATGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1599
QY 320 TGCGTCACAGCATGGCCAGAGATCATGTCAAAATATGCGATCTCTCAGAGGACAGAAATG 379
DB 1600 TGGCTCCAGACATGCGCAGAGAAACACACCAAGAGAGTCTTACAGCGAGAGCTGCCATC 1659
QY 380 TGCATGTCCAGTAAAGAGCCCTAGTGGAGAAATATCTCA 419
DB 1660 TGCGTGTGACAGCCCGCTGCGTTAGTGGCTGATATCTCA 1699

```
RESULT 8
US-09-736-960-1
; Sequence 1, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736, 960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160, 860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162, 498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170, 453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176, 195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182, 296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547, 276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196, 267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196, 460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196, 527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196, 528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687, 837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240, 503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240, 508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240, 539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240, 543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(6159)
; OTHER INFORMATION: human CLASP-5
US-09-736-960-1

Query Match      16.6%; Score 83.2; DB 10; Length 7215;
Best Local Similarity 50.5%; Pred. No. 1.5e-15;
Matches 202; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
```

```
DB      4669 AATAGCATCTTATATGACACAGTGAATGAGGAAATTCAGAGAAATTCAGATGCTT 4728
QY      260 GTGGACCTCCAGTACAGCCTGGCCAAATCTTATGCCAGCAGCCCGAGCTCAGGAAGACG 319
DB      4729 ATGCATCTCATGTACAGAAATTCAGAACTTACAGGCAATCTTGATCTGGGCTGACC 4788
QY      320 TGGCTGCAGACATGGCCAGAGATTCATCAAAATGGGATCTCTCAGAGGAGCAGCAATG 379
DB      4789 TGGCTGCAGAACATGGCAGAGAAACACACAGAAAGATGCTTACACGAGAGCTGCGATG 4848
QY      380 TGCTATGTCACCTAACAGCCCTAGTGGCAGAAATCTCA 419
DB      4849 TGGCTGTGACCGCGCTGCGTTAGTGGCTGATCTGA 4888
```

```
RESULT 9
US-09-796-692-5465
; Sequence 5465, Application US/09796692
; Patent No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algale, Paul A.
; APPLICANT: Menison, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796, 692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186, 126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190, 479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200, 545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200, 303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200, 779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200, 999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202, 084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206, 201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218, 950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222, 903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223, 416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223, 378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5465
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-5465

Query Match      15.8%; Score 79.4; DB 9; Length 427;
Best Local Similarity 53.4%; Pred. No. 6.3e-15;
Matches 167; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
```

```
QY      107 TCCCTGTCATCATCAACAACTGTGCCACAGTACCGGCTTATTAAGCACACAGCTTC 166
DB      10 TCCCTGAGGACAAATTTGGCTTATTCAGAAAGAGACACACCATGACAGATGCTCTTTT 69
QY      167 TCCCTGATGTGAAGGAGCTTAACCAAAAGATACGACGGTGAATGAGCCAGCCGACG 226
DB      70 CCCACCCAGGTGAGGAGACTTCTCTGTAATCTGATATGCATCTTATATGACAGTGA 129
```


OY	107	TCCCTGTCATCATCAACACTGTGCACAACATGACCAGCCTATTTAAGCACACAGCTTTC	166
Db	10	TCCTTGAGGACAATTTTGGCCTTATTACAGAAGAGGACACAGCATGCAGATGACTCTCTTT	69
OY	167	TCCTTGATGTGAAGAGCTTAACCAAAGGATACCGACGGTCTTAATGCCACGCCCCAG	226
Db	70	CCCACCCAGGTGGAGAACTTCTCTGTATATCGAATAGCATTTATATGACACACTGAAA	129
OY	227	ATGAAGGAGCAATGAAACGACGCCACAGATGCTGGTGGACCTTCACATACAGCCTGGCCAAA	286
Db	130	ATGAGGAGATTTTCAGAGAAATCTCTAGATGCTTAATGATCTCATGTATACAGAAATGGCAAG	189
OY	287	TCCTTATGCCACGACGCCCGAGCTCAAGAAAGCTGCTGCAGACATGGCCAGATTCAT	346
Db	190	AATTACCAAGGCAATCTCTGATCATGTCGGCTGACCTGCTCCAGAACATGG - CAGAGAACAC	248
OY	347	GTCAAAATTTGGGATCTCTCAAGGGCAGCAATGCTATGTCCAGATAACAGCCTTAGTG	406
Db	249	ACCAAGAAAGAACTGCTACACGAGGGCTGCATGCTCTGGTGCACGCCGCTGCTTAGTG	308
OY	407	GCAGAAATATCTCA	419
Db	309	GCTGAGTATCTGA	321


```

RESULT 11
US-09-736-968A-98
; Sequence 98, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-7 Transmembrane Protein
; FILE REFERENCE: 020054-000611US
; CURRENT APPLICATION NUMBER: US/09/736,968A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patentin Ver. 2.1

```

SEQ ID NO 98
LENGTH: 271
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: exon 116900-117170
US-09-736-968A-98

Query Match 11.5% Score 57.4; DB 9; Length 271;
Best Local Similarity 56.7%; Pred. No. 3e-08;
Matches 106; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 238 TGAGACGACCCAGAGATCGTGGACCTCCAGTACAGCCCGCCCAATTCCTATGCCAG 297
DB 4 TGAGTCCCTCTCCATACATCCCCCTCACCCTGGACTCCAGATTTGCCGGGCTTACCAGG 63
QY 298 CACGCCGAGCTCAGAGACGCTGCTCGACAGCATGGCCAGATCCATGTCAAAATGG 357
DB 64 CTCACCGGACCTTGGCTGACCTGGTTGCAGAACTGGCCGGGAGACACGGGACGTGG 123
QY 358 CGATCTCTAGAGGACGACATGTGCTATGTCCACGTAACGCCCTAGTGGCAGATATCT 417
DB 124 CAACCAAGCCGAGGCGCCGCACTGATGTGCACGCGCGCCCTGTGCTGATGACT 183
QY 418 CACACGG 424
DB 184 CGCCTG 190

RESULT 12

US-09-864-761-7634
Sequence 7634, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 7634
LENGTH: 512
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011472.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
US-09-864-761-7634

Query Match 11.5% Score 57.4; DB 10; Length 512;
Best Local Similarity 56.7%; Pred. No. 4e-08;
Matches 106; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 238 TGAGACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCCTGCGCCCAATTCCTATGCCAG 297
DB 169 TGAGTCCCTCTCCATACATCCCCCTCACCCTGGACTCCAGATTTGCCGGGCTTACCAGG 228
QY 298 CACGCCGAGCTCAGAGACGCTGCTCGACAGCATGGCCAGATCCATGTCAAAATGG 357
DB 229 CTCACCGGACCTTGGCTGACCTGGTTGCAGAACATGGCCGGGAGACCGGACGTGG 288
QY 358 CGATCTCTAGAGGACGACATGTGCTATGTCCACGTAACGCCCTAGTGGCAGATATCT 417
DB 289 CAACCAAGCCGAGGCGCCGCACTGATGTGCACGCGCGCCCTGTGCTGATGACT 348
QY 418 CACACGG 424
DB 349 CGCCTG 355

RESULT 13

US-09-864-761-24339
Sequence 24339, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24339
; LENGTH: 194
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011472.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
; OTHER INFORMATION: NT HIT: AB037816.1, EVALUE 2.00e-58
; OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUE 3.40e-01
; OTHER INFORMATION: EST_HUMAN HIT: BE531136.1, EVALUE 2.00e-87
US-09-864-761-24339
```

```
Query Match          10.0%; Score 50; DB 10; Length 194;
Best Local Similarity 58.9%; Pred. No. 4.8e-06;
Matches 86; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 279 TGGCCAATCTTATGCCAGACGCCGAGCTCAGGAAGACGTGGCTCGACAGCATGCCCA 338
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3 TTGCCGGGGGCTACGAGGCTCACGCGACCTTGGCTGACTGCTGTCAGAACATGGCCG 62

QY 339 GGATTCATGTCAAAATGGGATCTCTGACGAGGAGCAGCATGTGCTATGCCAGTAAAG 398
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 GGAAGCAGCGGGACGTGGGACACGCGAGGCCGCCCAAGTGTGTCACGCGGCCG 122

QY 399 CCCTAGTGGCAGATATCTCACAGG 424
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 CCTCGTGGCTGAGTACTGCGCCCTG 148

RESULT 14
US-09-736-960-76
; Sequence 76, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
```

```
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 266
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 13th exon
US-09-736-960-76
```

```
Query Match          9.7%; Score 48.8; DB 10; Length 266;
Best Local Similarity 58.1%; Pred. No. 1.3e-05;
Matches 86; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 272 TACACCTGGCCCAATTCCTATGCCAGCACGCCGAGCTCAGGAAGACGTGCTGCAGACG 331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 31 TTCACAATTTGCCAAGATTTACCAAGCATCTCTGATCTGGCGGTGACCTGCCTCAGAAC 90

QY 332 ATGGCCAGAGATCCATGTCAAAATGGCGATCTCTCAGAGCGCAGCAATGTCTATGTCCAC 391
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 91 ATGGCAGAGAACAACACCAAGAAAGTGTACACGAGGAGCTGCAATGTGCTGCTGCAC 150

QY 392 GTAAACGCCCTAGTGGCAGATATCTCA 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 151 GCCGCTGGCTTAGTGGCTGAGTATCTGA 178

RESULT 15
US-09-736-960-86
; Sequence 86, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
```

```
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 66686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ordered human genomic DNA at CLASP-5 locus
US-09-736-960-86
```

```
Query Match          9.7%; Score 48.8; DB 10; Length 66686;
Best Local Similarity 58.1%; Pred.No. 0.00016;
Matches 86; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 272 TACAGCCTGGCCAAATCTATGCGACGACGCCGAGCTCAGGAAGCGTGGCTCGACAGC 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33658 TTGAGAAATGGCCAAAGATTACCGAGCATCTCTGATCTCGGCTGACCTGCTCCGAGAC 33717

QY 332 ATGGCCAGGATCCATCTCAAAAAATGGGATCTCTCAGAGGACGACAAATGTCTATGTCCAC 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33718 ATGGCAGAGAAACACACAGAAAGTGTCTACACGAGCGTGCATGTGCCTGGTGCAC 33777

QY 392 GTACAGCCCTAGTGGCAGAAATATCTCA 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33778 GCCGCTGCTTAGTGGCTGAGTATCTGA 33805
```

Search completed: February 7, 2003, 09:14:41
Job time : 69.3549 secs

Tue Feb 11 17:06:31 2003

us-09-687-837-1_copy_2500_3000.rst

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 06:53:49 : Search time 1010.34 Seconds
(without alignments)
8030.908 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000

Perfect score: 501
Sequence: 1 ttacactggaagaagatcct.....gagagagcctcatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estoy:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estum:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	499.4	99.7	869	14 BQ434370	BQ434370 AGENCOURT
2	488.8	97.6	994	14 BQ067940	BQ067940 AGENCOURT
3	417.4	83.3	454	10 BE163028	BE163028 QV3-HT045
4	410.2	81.9	632	12 BG777371	BG777371 602664635
5	401.4	80.1	694	13 BE161316	BE161316 603309631
6	370	73.9	889	13 BE1695006	BE1695006 603345307

7	341.6	68.2	558	13 B070645	B070645 B070645
8	337.8	67.4	658	10 BB627919	BB627919 BB627919
9	330	65.9	503	10 AW909665	AW909665 u722d11.y
10	303.8	60.6	520	10 BE533728	BE533728 601232874
11	295	58.9	922	14 BQ424615	BQ424615 AGENCOURT
12	294.8	58.2	871	13 B1249092	B1249092 602992423
13	291.4	58.2	589	13 B1006821	B1006821 QV3-RT006
14	276.8	55.2	738	14 BQ179803	BQ179803 UI-M-EMO-
15	268.6	53.6	1042	13 BM456770	BM456770 AGENCOURT
16	263	52.5	615	12 BG087170	BG087170 H3136D04-
17	249.6	49.8	554	10 AM231163	AM231163 u040c09.y
18	243.4	48.6	600	12 BG806790	BG806790 2041-42.M
19	240.4	48.0	733	13 BG919189	BG919189 602818013
20	232	46.3	346	14 D58813	D58813 HUM512E07B
21	230.6	46.0	283	10 AM607903	AM607903 RCO-HT050
22	228.6	45.6	493	12 BG000343	BG000343 RCO-GN004
23	216	43.1	224	10 AM837402	AM837402 QV2-LT003
24	209.8	41.9	545	13 B1535482	B1535482 399059.MA
25	209	41.7	595	10 AW211391	AW211391 u080d10.y
26	207	41.3	1042	9 AL544595	AL544595 AL544595
27	206.8	41.3	349	9 AA384900	AA384900 EST98484
28	204.8	40.9	875	12 BE896788	BE896788 601437471
29	202	40.3	672	10 AW959926	AW959926 EST371997
30	195.2	39.0	327	9 AA384941	AA384941 EST98518
31	194.6	38.8	742	10 BE382198	BE382198 601271772
32	191.2	38.2	201	12 BF875977	BF875977 RCO-ET013
33	183.4	36.6	948	14 BQ213756	BQ213756 AGENCOURT
34	182.6	36.4	314	9 AA327686	AA327686 EST31082
35	173	34.5	870	14 BQ735599	BQ735599 AGENCOURT
36	172.4	34.4	731	10 AV383524	AV383524 AV383524
37	171.4	34.2	838	17 CNS0020E	AL097064 Drosophila
38	142.6	28.5	1101	17 CNS050F6	AL151548 Tetradodon
39	140.4	28.0	652	12 BP979009	BP979009 602147696
40	136.6	27.3	622	14 BQ390482	BQ390482 NISC_Mq13
41	136.6	27.3	773	12 BG664949	BG664949 DRABTD07
42	133	26.5	744	13 BG915748	BG915748 602814273
43	128.6	25.7	466	12 BG345674	BG345674 dd90f03.y
44	125	25.0	555	12 BF075158	BF075158 223976.MA
45	121	24.2	966	14 BM905401	BM905401 AGENCOURT

ALIGNMENTS

RESULT 1
BQ434370 869 bp mRNA linear EST 24-MAY-2002
LOCUS BQ434370
DEFINITION AGENCOURT_7827049 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:6152826
5' mRNA sequence.
ACCESSION BQ434370
VERSION BQ434370.1 GI:21173446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM13491 row: e column: 19
High quality sequence stop: 631.
Location/Qualifiers
1. .869

FEATURES

source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6152826"
/clone_id="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      223 a      232 c      236 g      178 t
ORIGIN
Query Match      99.7%; Score 499.4; DB 14; Length 869;
Best Local Similarity 99.8%; Pred. No. 1.2e-130;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTACACTGGAAGAAGTCTTTGTCGGACACATTTGCAAGTCATCATCTGTACGCCA 60
    |||||||
Db 117 TTACACTGGAAGAAGTCTTTGTCGGACACATTTGCAAGTCATCATCTGTACGCCA 176
QY 61 GGTGATAGCAGACGTTGTGGCATTTGGGGAACACAGATTCCAGCAGTCCCTGTCATCAT 120
    |||||||
Db 177 GGTGATAGCAGACGTTGTGGCATTTGGGGAACACAGATTCCAGCAGTCCCTGTCATCAT 236
QY 121 CAACAACCTGTGCCAACAGTGAACCGGCTTATTAGCACACACAGCTTCTCTGTGATGAA 180
    |||||||
Db 237 CAACAACCTGTGCCAACAGTGAACCGGCTTATTAGCACACACAGCTTCTCTGTGATGAA 296
QY 181 GGACTTAAACCAAAAGATACGACGAGTGCTATGGCCAGCCGACAGATGAAGAGCATGA 240
    |||||||
Db 297 GGACTTAAACCAAAAGATACGACGAGTGCTATGGCCAGCCGACAGATGAAGAGCATGA 356
QY 241 GAACGACCCAGAGATGCTGTGTGACCTTCAGTACAGCTTGCCAAATCTATGCCAGCAC 300
    |||||||
Db 357 GAACGACCCAGAGATGCTGTGTGACCTTCAGTACAGCTTGCCAAATCTATGCCAGCAC 416
QY 301 GCCCGAGCTCAGAGAGAGTGGCTCGACAGCATGGCCAGATCCATGTCTAAATATGGCGA 360
    |||||||
Db 417 GCCCGAGCTCAGAGAGAGTGGCTCGACAGCATGGCCAGATCCATGTCTAAATATGGCGA 476
QY 361 TCTCTCAGAGGACGATGCTGTATGTCACAGTACAGCCCTATGTGGCAGATATCTCAC 420
    |||||||
Db 477 TCTCTCAGAGGACGATGCTGTATGTCACAGTACAGCCCTATGTGGCAGATATCTCAC 536
QY 421 ACGGAAAGGCGTGTATTAGCAAGAGATGACCGCCTTCAGGAGTCATTACCCAAACATCGA 480
    |||||||
Db 537 ACGGAAAGGCGTGTATTAGCAAGAGATGACCGCCTTCAGGAGTCATTACCCAAACATCGA 596
QY 481 CGAGAGGCGCTCCATGATGA 501
    |||||||
Db 597 CGAGAGGCGCTCCATGATGA 617
RESULT 2
BO067940      994 bp      mRNA      linear      EST 02-APR-2002
LOCUS      AGENCOURT_6792185 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5768931
DEFINITION      5', mRNA sequence.
ACCESSION      BO067940
VERSION      BO067940.1 GI:19896986
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      1 (bases 1 to 994)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
```

```
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM12829, row: j column: 04
High quality sequence stop: 499.
FEATURES
    source
        location/Qualifiers
            1..994
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_image="5768931"
            /clone_id="NIH_MGC_121"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pCMV-Sport6; Site: 1; NotI;
            Site: 2; EcoRV (destroyed); RNA source anonymous pool of 3
            fetal brains, female age 20 weeks, female age 24 weeks,
            and male age 26 weeks. Library is oligo-dT primed and
            directionally cloned (EcoRV site is destroyed upon
            cloning). Average insert size 1.7 kb, insert size range
            0.7-3.5 kb. Library is normalized and enriched for
            full-length clones and was constructed by C. Gruber
            (Invitrogen). Research Genetics tracking code 017. Note:
            this is a NIH_MGC Library."
BASE COUNT      274 a      255 c      244 g      220 t      1 others
ORIGIN
Query Match      97.6%; Score 488.8; DB 14; Length 994;
Best Local Similarity 98.4%; Pred. No. 1.3e-127;
Matches 493; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 TTACACTGGAAGAAGTCTTTGTCGGACACATTTGCAAGTCATCATCTGTACGCCA 60
    |||||||
Db 187 TTACACTGGAAGAAGTCTTTGTCGGACACATTTGCAAGTCATCATCTGTACGCCA 246
QY 61 GGTGATAGCAGACGTTGTGGCATTTGGGGAACACAGATTCCAGCAGTCCCTGTCATCAT 120
    |||||||
Db 247 GGTGATAGCAGACGTTGTGGCATTTGGGGAACACAGATTCCAGCAGTCCCTGTCATCAT 306
QY 121 CAACAACCTGTGCCAACAGTGAACCGGCTTATTAGCACACAGCTTCTCTGTGATGAA 180
    |||||||
Db 307 CAACAACCTGTGCCAACAGTGAACCGGCTTATTAGCACACAGCTTCTCTGTGATGAA 366
QY 301 GCCCGAGCTCAGAGAGAGTGGCTCGACAGCATGGCCAGATCCATGTCTAAATATGGCGA 360
    |||||||
Db 487 GCCCGAGCTCAGAGAGAGTGGCTCGACAGCATGGCCAGATCCATGTCTAAATATGGCGA 546
QY 481 CGAGAGGCGCTCCATGATGA 501
    |||||||
Db 607 CGAGAGGCGCTCCATGATGA 501
    |||||||
QY 421 ACGGAAAGGCGTGTATTAGCAAGAGATGACCGCCTTCAGGAGTCATTACCCAAACATCGA 480
    |||||||
Db 607 ACGGAAAGGCGTGTATTAGCAAGAGATGACCGCCTTCAGGAGTCATTACCCAAACATCGA 666
QY 481 CGAGAGGCGCTCCATGATGA 501
    |||||||
Db 667 CGAGAGGCGCTCCATGATGA 687
RESULT 3
BE163028/c      454 bp      mRNA      linear      EST 21-JUN-2000
LOCUS      BE163028/c
DEFINITION      QV3-HT0457-070300-113-c03 HT0457 Homo sapiens cDNA, mRNA sequence.
```

ACCESSION BE163028
VERSION BE163028.1 GI:8625749
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 454)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?L1=6t2-QV3-HT0457-070300-113-c03&t3=2000-03-07&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 3
High quality sequence stop: 454.
Location/Qualifiers
FEATURES
source 1..454
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0457"
 /dev_stage="Adult"
 /note="Organ: head,neck: Vector: puc18; Site:1: Smal; Site:2: Smal: A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 88 a 112 c 124 g 130 t
ORIGIN
Query Match 83.3%; Score 417.4; DB 10; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.4e-107;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 10 AAGAGTCTTTGTCGCGACATTTGCAATCATATGTCAGCCAGCTGATAGC 69
 |||||
DB 454 AAGAGTCTTTGTCGCGACATTTGCAATCATATGTCAGCCAGCTGATAGC 395
OY 70 AGACGTTGTTGGCATTGGGAAACGATTCACAGAGTCCTGTCATCATCACTG 129
 |||||
DB 394 AGACGTTGTTGGCATTGGGAAACGATTCACAGAGTCCTGTCATCATCACTG 335
OY 130 TGGCAACAGTACCGGCTTATTAAAGCACACAGCTTCTCTGATGTGAAGACTTAC 189
 |||||
DB 334 TGGCAACAGTACCGGCTTATTAAAGCACACAGCTTCTCTGATGTGAAGACTTAC 275
OY 190 CAAAAGATACGACGAGTGTAAATGSCACCGCCAGATGAAGAGCATGAGAACGCC 249
 |||||
DB 274 CAAAAGATACGACGAGTGTAAATGSCACCGCCAGATGAAGAGCATGAGAACGCC 215
OY 250 AGAATGCTGTGGACTCCAGTACAGCCTGGCCAAATCTTAATGGCAGAGCCGAGCT 309
 |||||
DB 214 AGAATGCTGTGGACTCCAGTACAGCCTGGCCAAATCTTAATGGCAGAGCCGAGCT 155

OY 310 CAGGAAGACGTGGCTCGACAGCATGGCCAGATCCATGTCAAAATGCGATCTTCAGA 369
 |||||
DB 154 CAGGAAGACGTGGCTCGACAGCATGGCCAGATCCATGTCAAAATGCGATCTTCAGA 95
OY 370 GGCAGCAATGCTCTATGTGCCAGCTACAGCCCTAGTGGCCAAATATCTCCACAGGAAG 428
 |||||
DB 94 GGCAGCAATGCTCTATGTGCCAGCTACAGCCCTAGTGGCCAAATATCTCCACAGGAAG 36

RESULT 4
BG777371 632 bp mRNA linear EST 15-MAY-2001
BG777371
60266463551 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4604592 5',
mRNA sequence.
DEFINITION BG777371.1 GI:14047688
EST.
VERSION BG777371.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 632)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCID/DRP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM1651 row: e column: 09
High quality sequence stop: 620.
Location/Qualifiers
FEATURES
source 1..632
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4804592"
 /clone_lib="NIH_MGC_60"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site:1: SfiI (ggccgcctcgcc); Site:2: SfiI (ggccatagcgc) ; Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGCGGGGCGGCACATG-dT(30)BN-3 (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT 173 a 175 c 167 g 117 t
ORIGIN
Query Match 81.9%; Score 410.2; DB 12; Length 632;
Best Local Similarity 99.3%; Pred. No. 1.9e-105;
Matches 412; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 87 GGGAAACCAATTCACGACGTCCTGTCATCATCAACAACATGTCGAACGCCG 146
 |||||
DB 1 GGGAAACCAATTCACGACGTCCTGTCATCATCAACAACATGTCGAACGCCG 60
OY 147 TTATTAGCACACACACTTCTCTGATGTGAAGACTTAAACAAAGATAGCAGCG 206
 |||||
DB 61 TTATTAGCACACACACTTCTCTGATGTGAAGACTTAAACAAAGATAGCAGCG 120
OY 207 TGTCTATGCGCCCGCCAGATGAAGAGCATGAGAACGCCAGAGTCTGTGGACC 266
 |||||

Db 121 TGCATATGCGCACCGCCAGATGAAGAGCATGAGAACGACCCAGAGATGCTGTGCACC 180
Qy 267 TCCAGTACAGCCTGGCCAAATCTATGCCCACGCGCCGAGCTCAGGAAGAAGTGGCTCG 326
Db 181 TCCATACAGCCTGGCCAAATCTATGCCCACGCGCCGAGCTCAGGAAGAAGTGGCTCG 240
Qy 327 ACAGCATGGCCAGATTCATGTCAAAATGGCGATCTCTGAGAGGACGATGTGCTATG 386
Db 241 ACAGCATGGCCAGATTCATGTCAAAATGGCGATCTCTGAGAGGACGATGTGCTATG 300
Qy 387 TCCAGTACAGCCTGGCCAAATCTATGCCCACGCGCCGAGCTCAGGAAGAAGTGGCTCG 446
Db 301 TCCAGTACAGCCTGGCCAAATCTATGCCCACGCGCCGAGCTCAGGAAGAAGTGGCTCG 360
Qy 447 GCACCGCCTTCAGGTCATTTACCCCAATCGACGAGAGGCTCCATGATGGA 501
Db 361 GCACCGCCTTCAGGTCATTTACCCCAATCGACGAGAGGCTCCATGATGGA 415

RESULT 5
B1691316 694 bp mRNA linear EST 18-SEP-2001
LOCUS 603309631F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5345838 5',
DEFINITION mRNA sequence.
ACCESSION B1691316
VERSION B1691316.1 GI:15653945
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 694)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM11878 row: m column: 07
High quality sequence stop: 693.
Location/Qualifiers
FEATURES
source 1..694
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5345838"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 174 a 182 c 191 g 146 t 1 others
ORIGIN

Query Match 80.1%; Score 401.4; DB 13; Length 694;
Best Local Similarity 87.6%; Pred. No. 6, 2e-103;
Matches 438; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 2 TACACGTGAAGAAGATCTTGTCCGACACATTTCCAAAGTCAATCATCTGTACGCCAG 61
Db 146 TACACGGAAGAAGATCTTGTTCGAGCAGCACTTACAGTCACTATCTGTGACACCA 205
Qy 62 CTGATACAGACGTGTGTCATTTGGGGAACCAAGATTCAGACATGCTGTCCATCATC 121

Db 206 CTGATTCGACATGTGTTGGCATTTGGAGGACACAGATTCACAGAGTCTGTGCTATCATC 265
Qy 122 AACACTGTGCCACACAGTGAACGGCTTATTAAACACACACAGCTTCTCTCATATGGAAG 181
Db 266 AACACTGTGCCACACAGTGAACGGCTTATTAAACACACAGCTTCTCTCATATGGAAG 325
Qy 182 GACTTAAACCAAAAGATAGCAGGCTGCTAATGAGCCAGCCAGATGAAGAGCATGAG 241
Db 326 GATTTGACTAAGAGATCCGACAGATCTGTATGAGCCACAGCCAGATGAAGAGCATGAG 385
Qy 242 AACGACCAAGATGCTGTGAGCTCCAGTACAGGCTTGCCAAATCTATGCCAGCAGC 301
Db 386 AACGACCAAGATGCTGTGAGCTCCAGTACAGGCTTGCCAAATCTATGCCAGCAGC 445
Qy 302 CCGAGCTCAGGAAGACGTGCTGCAGACATGGCCAGATCCATGCTCAAAAATGGCGAT 361
Db 446 CTTGAGCTCAGGAAGACGTGCTGCAGACATGGCCAGATCCATGCTCAAAAATGGCGAT 505
Qy 362 CTCTCAGAGCAGCAATGTGCTATGTCACAGTAAACAGCCCTAGTGGCAGATATCTCACA 421
Db 506 CTCTCAGAGCAGCAGCAATGTGCTATGTCACAGTAAACAGCCCTAGTGGCAGATATCTCACA 565
Qy 422 CGGAAGCGCTGTTTACAGGAAGATGACCGCCTTACAGGTCATTAACCCCAACATCGAC 481
Db 566 CGGAAGCGCATGTTTCAGACAGGGGTGCACAGCCTTACAGGTCATTAACCCCAACATCGAT 625
Qy 482 GAGGAGCCTTCATGATGGA 501
Db 626 GAGGAGCCTTCATGATGGA 645

RESULT 6
B1695006 889 bp mRNA linear EST 18-SEP-2001
LOCUS 603345307F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5373121 5',
DEFINITION mRNA sequence.
ACCESSION B1695006
VERSION B1695006.1 GI:15657635
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 889)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM11949 row: n column: 02
High quality sequence stop: 652.
Location/Qualifiers
FEATURES
source 1..889
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5373121"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL3492 row: e column: 19

High quality sequence stop: 577.

Location/Qualifiers

1. 922

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6153210"

/clone_lib="NIH-MGC-67"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo

Average insert size 1.75 Kb. Library constructed by Life

Technologies."

BASE COUNT 251 a 237 c 249 g 185 t

ORIGIN

Query Match

Best Local Similarity 99.7%; Pred. No. 1.1e-72;

Matches 306; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 122 AACACGTGCGCAACAGTGCAGCGGCTTATTAGCAGACAGCTTCTCCTGTGATGTAAG 181

DB 1 AACACTGTGCC-ACAGTGCAGCGGCTTATTAGCAGACAGCTTCTCCTGTGATGTAAG 59

QY 182 GACTTAACCAAAAGGATGCGACAGGTTGCTATGGCCACCGCCAGATGAAGAGCATGAG 241

DB 60 GACTTAACCAAAAGGATGCGACAGGTTGCTATGGCCACCGCCAGATGAAGAGCATGAG 119

QY 242 AACGACCCAGAGATGCTGTGACCTCCAGTACAGCCTGCGCAAAATCCTATGCCAGCAG 301

DB 120 AACGACCCAGAGATGCTGTGACCTCCAGTACAGCCTGCGCAAAATCCTATGCCAGCAG 179

QY 302 CCCGAGCTCAGAGAGCTGCTGCAGCATGGCCAGAGATCCATGTCAAAATGGCGAT 361

DB 180 CCCGAGCTCAGAGAGCTGCTGCAGCATGGCCAGAGATCCATGTCAAAATGGCGAT 239

QY 362 CTCTCAGAGGCGAGCAATGTGCTATGTCACGCTAACAGCCCTAGTGGCAGAAATCTCACA 421

DB 240 CTCTCAGAGGCGAGCAATGTGCTATGTCACGCTAACAGCCCTAGTGGCAGAAATCTCACA 299

QY 422 CGGAAG 428

DB 300 CGGAAG 306

RESULT 12

BI249092 871 bp mRNA linear EST 17-JUL-2001

LOCUS BI249092/6 602992423p1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5148461 5'

DEFINITION

BI249092

VERSION

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgc.nci.nih.gov/

1 (bases 1 to 871)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL366 row: m column: 06

High quality sequence stop: 716.

Location/Qualifiers

1. 871

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:5148461"

/clone_lib="NCI_CGAP_Mam5"

/tissue_type="tumor, gross tissue"

/dev_stage="7 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo

library constructed by Life Technologies. Investigators

providing samples: Lohar Hennighausen/Robin Humphreys,

NIH"

BASE COUNT 175 a 245 c 225 g 226 t

ORIGIN

Query Match

Best Local Similarity 80.3%; Pred. No. 1.2e-72;

Matches 408; Conservative 0; Mismatches 92; Indels 8; Gaps 5;

QY 2 TACACTGGAAGAAGCTCTTGTCCGAGACATTTGCA--AGTCATCATCTGTGACG 58

DB 848 TACCAAGGAAGAAGCTCTTGTGCTGAGACGCGACAGTACAGGTCATCTCTGTGACG 789

QY 59 CAGC-TGATAGCAGAGCTTGTGGCAATGGGGAACACGATTCAGCACTCCCTGCTCAT 117

DB 788 CAACGTATTGCAAGATGTGTGCTCATTTGAGAGAACCAAGATTCAGCACTCCCTGCTCA 729

QY 118 CATCAACACTGTGCCAAGATGACCGGCTT--ATTAAACACA-CCAGCTTCTCTGTGA 174

DB 728 TCATCAACACTGTGCCAAGATGACCGGCTT--ATTAAACACA-CCAGCTTCTCTGTGA 669

QY 175 TGTGAAGACTTAACCAAAAGATATGACGAGGCTGCTAATGGCCACCGCCAGATGAAGA 234

DB 668 TGTGAAGACTTAACCAAAAGATATGACGAGGCTGCTAATGGCCACCGCCAGATGAAGA 609

QY 235 GCATGAGAACGCCAGAGATGCTGTGACCTCCAGTACAGGCTTGGCCAAATCTCATGCG 294

DB 608 GCATGAGAACGCCAGAGATGCTGTGACCTCCAGTACAGGCTTGGCCAAATCTCATGCG 549

QY 295 CAGCAGCCCGAGCTCAGAGAGAGCGGCTCGACAGCATGGCCAGAGATCCATGTCAAAAA 354

DB 548 CAGCAGCCCGAGCTCAGAGAGAGCGGCTCGACAGCATGGCCAGAGATCCATGTCAAAAA 489

QY 355 TGGCGATCTCTCAGAGGCGAGCAATGTGCTATGTCACGTAACAGCCCTAGTGGCAGATA 414

DB 488 TGGCGATCTCTCAGAGGCGAGCAATGTGCTATGTCACGTAACAGCCCTAGTGGCAGATA 429

QY 415 TCTCA-CAGGGAAGGCGGTGTTAGCAGAGATGACCCGCTTCAGGTCATTAACCCAA 473

DB 428 TCTCACCACGGAAGGCGGTGTTAGCAGAGATGACCCGCTTCAGGTCATTAACCCAA 369

QY 474 ACATCGACGAGAGCGCTCCATGATGA 501

DB 368 ACATCGACGAGAGCGCTCCATGATGA 341

RESULT 13

BI006821

LOCUS

DEFINITION QY3-RP0067-150101-538-a01 RP0067 Homo sapiens cDNA, mRNA sequence.

ACCESSION BI006821

VERSION BI006821.1 GI:14410879

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT	FEATURES	SOURCE
1	(bases 1 to 589)						
1	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsushima, A., Baid, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV3&t2=QV3-R00067-150101-538-a01&t3=2001-01-15&t4=1) Seq primer: puc 18 forward High quality sequence stop: 584.	location/Qualifiers 1..589 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="R00067" /dev_stage="Adult" /note="Organ: kidney-tumor; Vector: puc18; site_1: Sma1; site_2: Sma1; A mini-library was made by cloning products derived from ORESTRS PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
BASE COUNT	146 a	164 c	168 g	111 t			
ORIGIN							
Query Match	58.2%	Score 291.4	DB 13	Length 589			
Best Local Similarity	99.7%	Pred. No. 8.6e-72					
Matches 292	Conservative 0	Mismatches 1	Indels 0	Gaps 0			
QY	209	CTATATGGCCACCGCCCAAGTGAAGAGCATGAGACACCCCAAGATGCTGCTGACCTC	268				
Db	17	CCATATGGCCACCGCCCAAGTGAAGAGCATGAGACACCCCAAGATGCTGCTGAGACCTC	76				
QY	269	CAGTAGAGCCTTGAGCCAAATCCATATCCAGCAGCCGACGCTCAGAGAAAGAGTGGCTCGAC	328				
Db	77	CAGTAGAGCCTTGAGCCAAATCCATATCCATATCCAGCAGCCGACGCTCAGAGAAAGAGTGGCTCGAC	136				
QY	329	AGCATGGCCAGAGATTCATGTCAAAAATGGGATCTCTCAAGAGCAGCATGTGCTATGTC	388				
Db	137	AGCATGGCCAGAGATTCATGTCAAAAATGGGATCTCTCAAGAGCAGCATGTGCTATGTC	196				
QY	389	CAGCTAACAGCCTTGATGGCAGAAATATCTCACAGGAAAGGCGTGTATAGACAAGATGC	448				
Db	197	CAGCTAACAGCCTTGATGGCAGAAATATCTCACAGGAAAGGCGTGTATAGACAAGATGC	256				
QY	449	ACCGCCTTACAGGCTCATTAACCCCAACATGACAGAGAGGCGCTTCATGATGGA	501				
Db	257	ACCGCCTTACAGGCTCATTAACCCCAACATGACAGAGAGGCGCTTCATGATGGA	309				
RESULT 14	B0179803	738 bp	mRNA	linear	EST 30-APR-2002		
LOCUS	B0179803	UT-M-EMD-bow-d-19-0-UT-1	NIH-BMAP-EMD	Mus musculus	cDNA clone		
DEFINITION	B0179803	IMAGE:5703138	5', mRNA				
ACCESSION	B0179803.1	GI:20355295					
VERSION							

	KEYWORDS	EST.
	SOURCE	house mouse.
	ORGANISM	Mus musculus
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	AUTHORS	Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
	TITLE	1 (bases 1 to 738)
	JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. James Lin, University of Iowa CDNA library preparation: Dr. M. Bento Soares, University of Iowa CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIN at: http://image.lim.nih.gov This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
	FEATURES	Seq primer: pyX-5.
	SOURCE	Location/Qualifiers
		1..738
		/organism="Mus musculus"
		/strain="C57BL/6"
		/db_xref="taxon:10090"
		/clone_1IMAGE:5703138"
		/clone_1lib="NIH_BMAP_EMO"
		/tissue_type="whole brain"
		/dev_stage="embryo 15.5 dpc"
		/lab_host="DHI0B (T1 phage resistant)"
		/note="Organ: brain; Vector: pyX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail , is GTGCGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institutes of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."
	BASE COUNT	182 a 209 c 173 g 174 t
	ORIGIN	
	Query Match	55.2%; Score 276.8; DB 14; Length 738:
	Best Local Similarity	87.8%; Pred. No. 1.4e-67:
	Matches 302; Conservative	0; Mismatches 42; Indels 0; Gaps 0;
Yy	2 TACACTGGAAAGATCTCTTTGTCCGACACATTGGCAAGTCATCATATCTGTACGCCAG	61
Db	395 TACACAGGAAGAAGATCTTTGTGGAGCGACATTACAGTCATCATCTGTCAACCNA	454
Oy	62 CTGATACACACACTTTGGCATTTGGGAAACCGAGTTCCAGACAGTCCCGTCATCANC	121
Db	455 CTGATTCGACATGTGTTGGCATTTGSAGAACCGAGATTCCAGCAGTCTTGTCTATCANC	514
Oy	122 AACCACTGTGCCAACAGTAGCACGCGCTTTTAAGCACACACAGCTTCTCTCTGATGTGAAG	181
Db	515 AACCAACTGTCCCAACGACCGACCGATCATCAAGCACACACAGCTTTCTCTCTGATGTGAAG	574
Oy	182 GACTTAACCAAAGATAGTCAGTCAGTGTATATGGCCACCGCCAGATGAAGAGCATGAG	241
Db	575 GATTTCGCTAAGAGATCCGCACAATCTCTGATGGCCACAGCCAGATGAAGAGACAGAG	634
Oy	242 AAGCACCAGACAGTCTGTGTGAGACCTCCAGTAPACAGCTGGCCAAATCTTATGCCACAGC	301
Db	635 AACGACCCGAGATCTGTGTGAGACCTCCAGTAPACAGCTGGCTAAGTCTTACAGCCAGACAC	694

Oy 302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATTTGCCAGATCA 345
||| |||||
Db 695 CCTGAGACTCAGGAAGACGTGGCTAGACAGTATTGGCGAGGATTCA 738

RESULT 15

LOCUS	1042 bp	mRNA	linear	EST 05-FEB-2002
DEFINITION	BM456770	AGENCOURT_6404261	NIH_MGC_92 Homo sapiens	CDNA clone IMAGE:5583671
	5', mRNA sequence.			

ORIGINISM	REFERENCE
Homio, superciliata	Ekström, Marmosa; Chorodita; Craniata; Vertebrata; Euteleostomi;
Ekström, Marmosa	Mammalia; Mammalia; Primates; Cetartiodactyla; Homiidae; Homo.
Mammalia; Mammalia	1 (bases 1 to 1042)
NH MGC	http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1998)
Contact: Robert Strussberg, Ph.D.	

	a	c	g	t	others
BASE COUNT	344	190	221	286	1
ORIGIN					

Query Match	53.6%;	Score 268.6;	DB 13,	Length 1042;
Best Local Similarity	72.3%;	Pred. No. 3.5e-65;		
Matches 349;	Conservative	0;	Mismatches 134;	Indels 0;
				Gaps 0.

Qy	19	CTTTGCCGACCACTTTCGASTGCTCATCTATCTGTGCACCCAGCTGATAGCAGACTTGT	78
Db	2	CTTTTGGAGCACATCTACAGATTAAATTGCTGTGAACCAACGTATAGCTGATGTAGC	61
Qy	79	TGGCATTGGGGAAACCAAGATTCAGCAGTCCCTGTCCATCATCAACTGTGCCAACG	138
Db	62	ACTAAGCGGAGGATCAAGATTTCAGAGTCTTATTCATTATCAATTAATTGGCAATAG	121
Qy	139	TGACCGGCTTATTAAGACACACCGCTCTCTCTGTGATGAAGGACTTAAACCAAAAGAT	198
Db	122	TGACAGACCTATGAGGCAACTGCTTTCCCGAGAAAGTCAAAAGACTTGACCAAGAGAT	181
Qy	199	ACGCAGGTGCTTAATGGCCACCGCCCAAGATGAAGAGCATGAGAACGACCAGAGATGCT	258
Db	182	CCGCACCTGTTCTTATGGCCACTCTCCCAAAATGAAGAGGCAATGAGAAAGACCTCGAAATGCT	241
Qy	259	GGTGACTCTCAGTACAGCCTGGCCAAATCTATGCGACAGACGCCGAGACTCAGGAAGAC	318
Db	242	AATGATGCTCCAGTATAGCTTATAGCAAGGCTCTATGCAAGAGACCCAGACTCAGGAAAC	301
Qy	319	GTGGCTGCACAGCATGGCCAGAGTCCATGTGCAAAAATGGCGATCTCTCAGAGGCAACCAT	378

Db	302	CTGGCTTGAATACCATGCGCCAGATTCATGTAAAAATGGAGATTTTTCAGAGCGCTGGAT	361
Qy	379	GTGGTATGTCACGTAACAGCCCTAGTGGCAGATACTTCACACGAAAAGGCGTGTAG	438
Db	362	GTGTATAGTCCATGTAGCAGCTCTTAGTTGCAGAGTTCTTCATCGAAAAATATATTTCC	421
Qy	439	ACAAGATGCACCGCCTTCAGGGTCATTACCCCAACATCGACGAGAGCGCTCCATGAT	498
Db	422	TAACGATTGTCAGGCTTCAGAAATTACTCCCAATATGATGAAGAAGAGCAATGAA	481
Qy	499	GGA 501	
Db	482	AGA 484	

Search completed: February 7, 2003, 08:49:35
Job time : 1018.34 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:39:15 ; Search time 996.663 Seconds

(without alignments)
14629.322 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300

Perfect score: 501

Sequence: 1 caattgtggaagcttcg9.....ttgtcttaacaagaagtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 1451402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ha: *
2: gb_hg: *
3: gb_hn: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pi: *
9: gb_pr: *
10: gb_ro: *
11: gb_sy: *
12: gb_un: *
13: gb_un: *
14: gb_vt: *
15: gb_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_ov: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vt: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rnd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	7506	6	AX255048
2	490	97.8	3899	6	AX058220
3	471	94.0	7522	9	AF527605
4	471	94.0	7545	2	AB028981
5	365	72.9	155198	2	AL357553
6	365	72.9	163316	9	AL161420
7	328.4	65.5	2998	10	BC009134
8	278.8	55.6	3227	10	RNTRG
9	220.6	44.0	182640	2	AC126253
10	220.6	44.0	209372	2	AC126033
11	206.2	41.2	175281	2	AC109966
12	175.6	35.0	192825	2	AC119357
13	79.4	15.8	2768	9	AB056820
14	78.2	15.6	2036	9	AK001253
15	77.6	15.5	16884	2	BC015018
16	77.6	15.5	16884	2	AC023985
17	77.6	15.5	171811	9	AC011739
18	74.4	14.9	51808	5	AC104683
19	71.2	14.2	162378	2	AC105541
20	71.2	14.2	174041	2	AL772340
21	67.8	13.5	2422	9	AK000227
22	67.4	13.5	2299	9	AK054649
23	67.4	13.5	2610	6	AK058192
24	67.4	13.5	3152	9	AK056684
25	67.4	13.5	3472	6	AX173120
26	67.4	13.5	4391	6	AX173028
27	67.4	13.5	4393	6	AX173118
28	67.4	13.5	6454	6	AX173022
29	65	13.0	153026	2	AC094791
30	63.4	12.7	103906	2	AL672038
31	63.4	12.7	181127	10	AL672038
32	56.2	11.2	139887	2	CNS01RGX
33	44.4	8.9	200100	2	AC115789
34	42.4	8.5	146383	2	AC116367
35	41.8	8.3	244	6	AX174652
36	41.8	8.3	1476	6	AX399260
37	41.8	8.3	1605	6	AX174666
38	41.8	8.3	1669	9	HSMB02936
39	41.8	8.3	2014	6	AX174662
40	41.8	8.3	2904	6	BC019102
41	41.8	8.3	4026	6	AX174571
42	41.8	8.3	4027	6	AX174664
43	41.8	8.3	4577	9	AK024436
44	41.8	8.3	5564	9	AK074081
45	41.8	8.3	6422	9	AK090429

ALIGNMENTS

RESULT 1
AX255048
LOCUS AX255048
DEFINITION Sequence 7 from Patent WO0170808.
ACCESSION AX255048
VERSION AX255048.1 GI:16074541
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 7506)
AUTHORS Rastelli, L.K. and Gerritsen, M.
TITLE Angiogenesis-associated proteins, and nucleic acids encoding the same

JOURNAL Patent: WO 0170808-A 7 27-SEP-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES Location/Qualifiers
source 1..7506
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2147 a 1689 c 1764 g 1906 t
ORIGIN

Query Match 100.0%; Score 501; DB 6; Length 7506;
Best Local Similarity 100.0%; Pred. No. 2,8e-133;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTTGTGAGAGCTTGGCTCAAGCCTTACGGGTAAAGCAAGCTGTCTATTAAAGAGAC 60
DB 6106 CAATTTGTGAGAGCTTGGCTCAAGCCTTACGGGTAAAGCAAGCTGTCTATTAAAGAGAC 6165
QY 61 CAGCTGATGATGAGAGAAATGAAGCAACTACAGGAATGCGAAGAGGCTTCT 120
DB 6166 CAGCTGATGATGAGAGAAATGAAGCAACTACAGGAATGCGAAGAGGCTTCT 6225
QY 121 GAAATCATGATGAGAGAGATCTGCCCCCTGGAGAGAGAGAGAGAGCTTTACGGAATTC 180
DB 6226 GAAATCATGATGAGAGAGATCTGCCCCCTGGAGAGAGAGAGAGAGCTTTACGGAATTC 6285
QY 181 CTTGCATCTTCAACGCCATCAGTGGGAGCTCCAAAGCAAGCAATGGTTACGGATGAC 240
DB 6286 CTTGCATCTTCAACGCCATCAGTGGGAGCTCCAAAGCAAGCAATGGTTACGGATGAC 6345
QY 241 AGCTGCTCTTGGCTGCTGATTTACATCTCATGCGCCGTGTGTGGGAGCTTCTTGTCA 300
DB 6346 AGCTGCTCTTGGCTGCTGATTTACATCTCATGCGCCGTGTGTGGGAGCTTCTTGTCA 6405
QY 301 TTTCGAACACGAGAGTCTTCCAAAGCCATACGAGGAGAGAGAGAGAGAGAGAGAG 360
DB 6406 TTTCGAACACGAGAGTCTTCCAAAGCCATACGAGGAGAGAGAGAGAGAGAGAGAG 6465
QY 361 CAAGGGAG 420
DB 6466 CAAGGGAG 6525
QY 421 AGGAGTGTAGAGAGTGCATATTTTAAATCTCACTGGCAATATTCAGAGTTTC 480
DB 6526 AGGAGTGTAGAGAGTGCATATTTTAAATCTCACTGGCAATATTCAGAGTTTC 6585
QY 481 ATTGTGCTTTACAAAGGTGT 501
DB 6586 ATTGTGCTTTACAAAGGTGT 6606

RESULT 2
AX058220 3899 bp DNA linear PAT 17-JAN-2001
LOCUS AX058220
DEFINITION Sequence 90 from Patent WO0077040.
ACCESSION AX058220
VERSION AX058220.1 GI:12310721
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3899)
AUTHORS Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Bandman,O., Baughn,M.R.,
Atimel,Y., Yang,J., Reddy,R. and Lu,D.A.
TITLE Human intracellular signaling molecules
JOURNAL Patent: WO 0077040-A 90 21-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..3899
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1078 a 929 c 949 g 943 t
ORIGIN

Query Match 97.8%; Score 490; DB 6; Length 3899;
Best Local Similarity 99.8%; Pred. No. 3.9e-130;
Matches 501; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CAATTTGTGAGAGCTTGGCTCAAGCCTTACGGGTAAAGCAAGCTGTCTATTAAAGAGAC 60
DB 3345 CAATTTGTGAGAGCTTGGCTCAAGCCTTACGGGTAAAGCAAGCTGTCTATTAAAGAGAC 3404
QY 61 CAGCTGATGATGAGAGAAATGAAGCAACTACAGGAATGCGAAGAGGCTTCT 120
DB 3405 CAGCTGATGATGAGAGAAATGAAGCAACTACAGGAATGCGAAGAGGCTTCT 3464
QY 121 GAAATCATGATGAGAGAGATCTGCCCCCTGGAGAGAGAGAGAGAGCTTTACGGAATTC 179
DB 3465 GAAATCATGATGAGAGAGATCTGCCCCCTGGAGAGAGAGAGAGAGCTTTACGGAATTC 3524
QY 180 CTTGCATCTTCAACGCCATCAGTGGGAGCTCCAAAGCAAGCAATGGTTACGGATGAC 239
DB 3525 CTTGCATCTTCAACGCCATCAGTGGGAGCTCCAAAGCAAGCAATGGTTACGGATGAC 3584
QY 240 CAGCTGCTTGGCTGCTGATTTACATCTCATGCGCCGTGTGTGGGAGCTTCTTGTCTC 299
DB 3585 CAGCTGCTTGGCTGCTGATTTACATCTCATGCGCCGTGTGTGGGAGCTTCTTGTCTC 3644
QY 300 ATTTCGAACACGAGAGTCTTCCAAAGCCATACGAGGAGAGAGAGAGAGAGAGAGAG 359
DB 3645 ATTTCGAACACGAGAGTCTTCCAAAGCCATACGAGGAGAGAGAGAGAGAGAGAGAG 3704
QY 360 CCAAGGGAG 419
DB 3705 CCAAGGGAG 420
QY 420 TAGGAGTGTAGAGAGTGCATATTTTAAATCTCACTGGCAATATTCAGAGTTTC 479
DB 3765 TAGGAGTGTAGAGAGTGCATATTTTAAATCTCACTGGCAATATTCAGAGTTTC 4824
QY 480 CATTGTGCTTTACAAAGGTGT 501
DB 3825 CATTGTGCTTTACAAAGGTGT 3846

RESULT 3
AF527605 7522 bp mRNA linear PRI 01-AUG-2002
LOCUS AF527605
DEFINITION Homo sapiens zizimin1 mRNA, complete cds.
ACCESSION AF527605
VERSION AF527605.1 GI:22038158
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 7522)
AUTHORS Meller,N., Iranl-Tehrani,M., Kiosses,W.B., Del Pozo,M.A. and
Schwartz,M.A.
TITLE zizimin1, a novel Cdc42 activator, reveals new guanine nucleotide
exchange-exchange factor domain for rho proteins
JOURNAL Nat. Cell Biol. (2002) In press
AUTHORS Meller,N. and Schwartz,M.A.
TITLE Direct Submision
JOURNAL Submitted (05-JUL-2002) Cell Biology, The Scripps Research
Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES Location/Qualifiers
source 1..7522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
56..6265
/note="guanine nucleotide exchange factor; Cdc42
activator"
/codon_start=1


```
/product="z12im1n1"
/protein_id="AAM90306.1"
/db_xref="GI:22038159"
/translation="MSQPLPASPASAEKRTKRTALSKPGTAELRQSVSEVVGSVLLA
KPKLEPDIENVIYOKTQILINDCLREMLPEYDQDQALIRQGRICSTVPAKRE
EASGLPVEKICKYNSDMHLVYKYEYSEGFQPLPKLVHLPVHYVEVEVD
KDEDAISCSQKGGITTKHGM.YKGMNSAISTVMSRFRPFH.II.DGDSYLNPFK
DEKISKEKGSIFLDSCGVYONNKVRFPAELKQKDSIFLLAADSVEBEMETTL
NKIIQLNEEAMQEKRNDSHDEDSQSLGSGSDSYLPELKLKSAEALIKKSES
RKLEFLDPDAQKIDFSSAEPEVKSFEKPKRILVKNCDNISPALMGSGOSPVLKI
NPEPFTVLSLDIKYNNKISADFHVDLNFHSVQM.LATSPALMGSGOSPVLKI
LHEAMQYKPGISVTCPPHPIFLVATIEVLOSILHCAEPYNSDSSSKVAOKVL
KNAKACORLQGYRMPFAMAARTL.FKDAISGLDKARPAALTYRQDSNKLSDMLKLL
ADFRPEKMAKLPVLGLNDITIDVSSDPFNYSSTIPPKQETGCKPTITVEE
FVPCIPKHTOPTITNHLVYPKYLKYDSQSKAKANAI.CLEFKDSDESDQPLK
CIYRPGGVPTFRSAFAVLHHHONPEFYDKIELPQLEKHLLLTFFHVSQDNS
SKGSTKRQDVETQVGYSMPLKDGKRVTSQOH.PVANSPLSGALGOELGMGRHYG
PEIKRWGDKGKPLKISTHLVSTVYQDQHLNFPQYCOKSTSGAOLGELVYKYLKL
HAMEGHVMAPIPTLNLFRVLPTRATQEVAVNVRV.IIHVVAQCHEGLESRLSY
VAYATKAEPYVASEKTYHBEELKSMITLILFSAFLSNKLKTSFPEVYLKMSA
OHLIENSKVKLRNORFPASVYHAYEVVNM.LMPHTOKFDNPEPASNANHS.LAVET
KRCFTFMDRGVEFKQINNYISCFAPGDKTLFEYFEFLRVVCNEHYIPLNLPFG
KGIQRYODLQDYSLDEPCFNHFLVGLLREVGTAQOEREVLI.ISVLKNLIK
HSEDVYASRSHQARIATLYLPLFGLLIENORIVRQVSPPVAGWTVDESLALP
AVNPLVTPKQGSTLDNSLKHLDLGAISGLASPTTSTPNINSVRNADRSGLSTDSG
NVLPEKNSKSLDKHOOSTLGSVVRCDLIDQSEIKSLMLCKLITLKSQDALL
TYMNAKSTSELMDEFTISEVCLHOFQYMGKRIARTGMHMLDQGLSDLSLTFNHS
YGHSDADVHOSLELANIATEVCLTALDTLSLFTLAFNOLADGHNP.LMKVYDV
LGFLOHOSETALKNVFALRSLIYKFPSTFEGADMACALCYEILCCNSKLSIR
TASOLLYLNNNDYQKGSFVRTHLOV.IISVQSLADVYIGCTGFEOSLSIINN
CANSPRLIKHTSPSSDYDLPKRTVLTQMTQOMEHEHNDPMDIOLYLSAKVAST
PELRKWTDSMARIRHKNKNDLSEAMCYVHTALVATVITREKAVOMPEPLPHSHA
CLRSKGVFROGCTAFRVTPTNIDEEASMMEDVOMQHEMEDVLMELREOCALGSA
KABRELADIYKLIPIYKRRDEREALH.LHRAYSKTEVYKSGRGLLGYFR
VAFEGCAQYQFTSETDVEGFEDDEGEYLYKPEKLTPLSEIQRILKYIATFEM
ENWKMIDQSGKVPKDLDSKAYIOVTHV.IPFEDKELOERTEPERSNIRFPEM
PTQMGKRGVGEBOCKRRTILTAIHCFPYVKRK.IPVMOCHTDLNLEVALIDEMSK
VAELRQCSAEVDMKIDOKIQGSSVQVNNAPLARAFLPDINTRYDPNKKLL
KEVFRQFVACQDALAVNERLLKEQDLEIYQEBKANYRMAKELSEIMHEQLG"
```

BASE COUNT 2146 a 1694 c 1774 g 1908 t
ORIGIN

Query Match 94.0%; Score 471; DB 9; Length 7522;

Best Local Similarity 98.2%; Pired. No. 1,2e-124;

Matches 500; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

```
QY 1 CAATTGTGGAAGCTTGGCTCAAGCCTTAGCGGTAACGCAAGCTGTGATTAAAGAAGAC 60
|||||
Db 6119 CAATTGTGGAAGCTTGGCTCAAGCCTTAGCGGTAACGCAAGCTGTGATTAAAGAAGAC 6178
|||||
QY 61 CAGCTCAGATATCAGGAAGAATGAAAGCCACTACAGGGAATGGCGAAGAGCTTTCT 120
|||||
Db 6179 CAGCTCAGATATCAGGAAGAATGAAAGCCACTACAGGGAATGGCGAAGAGCTTTCT 6238
|||||
QY 121 GAATCATGCATGCA-----GATCTGCCCCCTGGAGGAGAGACAGCGCTTAC 172
|||||
Db 6239 GAATCATGCATGCAAGCTGGAGATCTGCCCTGGAGGAGAGACAGCGCTTAC 6298
|||||
QY 173 CGAATTCGCTTACATCTTCAACGCCATCAGTGGAGCTCCAAACAGACACATGGTTACG 232
|||||
Db 6299 CGAATTCGCTTACATCTTCAACGCCATCAGTGGAGCTCCAAACAGACACATGGTTACG 6358
|||||
QY 233 GGATGACACAGCTCTCTTGGCTGTGATTACATCTCATGGCCCGTGTGTGGGACTTG 292
|||||
Db 6359 GGATGACACAGCTCTCTTGGCTGTGATTACATCTCATGGCCCGTGTGTGGGACTTG 6418
|||||
QY 293 CTTTGTCTATTGGCAACTCAGAGATGCTTCCAAAGCCAAATCAGTGGGAGACCGAGACA 352
|||||
Db 6419 CTTTGTCTATTGGCAACTCAGAGATGCTTCCAAAGCCAAATCAGTGGGAGACCGAGACA 6478
|||||
QY 353 GGGAGAGCACAAGGGAGGAGAGAAATTAAGAACACGTTATTCTTAACAGA 412
|||||
Db 6479 GGGAGAGCACAAGGGAGGAGAGAAATTAAGAACACGTTATTCTTAACAGA 6537
|||||
```

```
QY 413 CTTCTATAGAGCTTGTAGAGAGTGCACAPATTTTAAATCTCAGTCATATTTCA 472
|||||
Db 6538 CTTCTATAGAGCTTGTAGAGAGTGCACAPATTTTAAATCTCAGTCATATTTCA 6597
|||||
QY 473 AAGTTTCATTTGCTCTTAAACAAGGTGT 501
|||||
Db 6598 AAGTTTCATTTGCTCTTAAACAAGGTGT 6626
|||||
RESULT 4
AB028981
LOCUS 7545 bp mRNA linear PRI 10-MAY-2002
DEFINITION Homo sapiens mRNA for KIAA1058 protein, partial cds.
ACCESSION AB028981
VERSION AB028981.2 GI:20521745
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Kikuno,R., Nagase,T., Ishikawa,K., Hirose,M., Miyajima,N.,
Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XIV. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 6 (3), 197-205 (1999)
MEDLINE 99397452
PUBMED 10470851
REFERENCE 2 (bases 1 to 7545)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@nfoekazusa.or.jp, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
COMMENT On May 9, 2002 this sequence version replaced gi:5689452.
FEATURES
source
```

```
gene
CDS
1..7545
/gene="KIAA1058"
<1..6288
/gene="KIAA1058"
/codon_start=1
/product="KIAA1058 protein"
/protein_id="BAAB3010.2"
/db_xref="GI:20521746"
/translation="IACGLARPGTRSCSCRLSDLRFPQPMQADKRTSRSYKEL
VIESPLQKDAQGEVFAESPVPAPKRELFYEDVENVIVOKTQILINDCLREMLLF
PYDPQQTALIRQGRYICSTYPAKAEPAQSLFYTECIIKYNNSMHLVYKYEYSGE
FRLDPKAVKLDLDPVHYEVDEEDKDEDAISLSQKGGITTKHGM.YKGMNSAISTV
TMSRFRPFH.II.DGDSYLNPFKDEKISKERGSTFLDSCGVYONNKVRFPAE
LKMQKSSYLLAADSEVMEEMWITLNL.IQLNEEAMQEKRNDSHDEDSQSLGSG
SDSYLPELAKSAREAEIKLSSEKRLFLDPDAQKIDFSSAEPEVKSFEKPKR
RLVYKCNLSFNLCQVANEQGTINEPFTVLSLFDIKYNNKISADFHVDLNFHS
VRQMATSPALMGSGOSPVLVIGIIEHAMOYKPGIISVTCPPHPIFLVATIEV
LOGSTHCAEPYKSSDSKVAQYLYKAKACORLQGYRMPFAMAARTL.FKDSGSL
DKNARFSAIYRQDSNKLSDMLKLDLDFKPEKMAKLPVLGLNDITIDVSSDPFN
VYNSYIPTKQETCKPTIFEEVEEYPCIPKHTOPTITNHLVYVYKYLKYDSQ
SFAKRNAIACIEFKDSDESDQPLKCIYRPGGVPTFRSAFAVLHHHONPEFYDEI
KIELPQJLHEKHLLLTFFHVSQDNSKSGSKRQDVVETQVGYSMPLKLDGKRVVTS
```

QHIPVSNLPSCGYLQYBLGKGRHYEPKRWGDKRPLKISTHLSVTVYQDQHLN
FPOYCKTESGAOLAGNELVYLKSLHMECHWIAIPLTLLNLFVLIRAOEVA
VAVTRIIHVNOCHEEGELSHRSYVAYKAPRYASEKTYELLETSMITLKP
SADPLSNKLKYSNEFFDVLKISMAOHLIENSVKRLRPFASVHHNVEVNL
MHIIQKPFDPNDEASKNNHNSHAFIKRCPFMNREGVFNQINNYICAPGDKTIF
ETKFEELRVNCHENTIPLANDPFPKGRIRORYDOLDISLTDEFFRNHFLVLLR
EVGTALOEERREVLIAISVKNLKHSPDRVIRASRHOATATLYLPTELLENVO
RINRVSPFPVNAAGTVKDESLAPANPLVTPOKSTLONSJHKLDAISGIAP
YTTSTPINSVBNADSRGLSTDSGNSLPERNSEKSLDHQOSTLNSVYCKD
LDOSEIKSLMCFEYLKMSMDALFTYWNKASTSELDPEFTISEVCLHQFOYMKRY
IARTGMMHARLOGLSDLSLNFNSHSGADVHOSLEPANTATVCLADLTSL
FTLAKNOLADHGNHKNPKYFVYLCFLOKHOSETAKRPFALSLIYKPESTFY
EGRADMCALCEIILKCCNSKISIRTSASOLITFLMNNNDYTGKSPRTHOYII
SVSOLADVIGIGTRFOOSLSIINNCAINSRPLIKHSESDVADLKRITVLMATA
OKMEHNDPEMLVDLOYSLAKSYASTPELRKRWDSMARIHKNGLSEAMCYVHT
ALVAEELTRKEAVOWEPPLPHSHSACLRSGRGVFRGACAFVITPNDIEASME
DYMODVHNEDEVIMELLECADLMKARELIADIKLIPIEKERDPERIAHLY
DILHRAYSVTEVMSHGRILGTFRVAFPGOAOYOFTDESETVEGFEDEDEKEYI
YKPKPLPSLSEIORSILKISPKFGSEVKNMIOGSKVNPRLDLSKAYIIOVTHVPIY
FDEKELOEKRTTEFERSHNRFRMFEMPTOTGRGKRGVGEDECKRRTITLTAHVPIYK
KRIPVMTQHTDNLPIEVAIDEMSKVAELNOLCSSAVEDMIKIQLKQGSVSOVA
GPLAVARAFLDNTNTRKYPDNKVKLLKEVROFVACGQALAVNERLIKEDOLEYOE
MKANYREMAKELSEIMHEDLG"

BASE COUNT 2175 a 1687 c 1753 g 1930 t

ORIGIN

Query Match 94.0%; Score 471; DB 9; Length 7545;
Best Local Similarity 98.2%; Pred. No. 1.2e-124;
Matches 500; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

QY 1 CAATTGTGAGAGCTTCGCTCAAGCCTTAGCGGTAAACGATCTGATTAAGAAGAC 60
DB 6142 CAATTGTGAGAGCTTCGCTCAAGCCTTAGCGGTAAACGATCTGATTAAGAAGAC 6201
QY 61 CAGCTCGAGTATCAGAGAAATGAAAGCCACTACAGGAAATGCGGAGAGCTTTCT 120
DB 6202 CAGCTCGAGTATCAGAGAAATGAAAGCCACTACAGGAAATGCGGAGAGCTTTCT 6261
QY 121 GAATTCATGCTGAGCA-----GATCGCCCTCGAGAGAGAGACGACGCTTAC 172
DB 6262 GAATTCATGCTGAGCAAGCTGGATGATCTGCCCTCGAGAGAGAGACGACGCTTAC 6321
QY 173 CGAATTCCTTACATCTTCAAGCCATCAGTGGAGCTCCCAACAGACATGTTACG 232
DB 6322 CGAATTCCTTACATCTTCAAGCCATCAGTGGAGCTCCCAACAGACATGTTACG 6381
QY 233 GGATGACAGCTGCTTCGGTGGTGGATTACATCATGCGCCGTGTGGGACTTG 292
DB 6382 GGATGACAGCTGCTTCGGTGGTGGATTACATCATGCGCCGTGTGGGACTTG 6441
QY 293 CTTTGTCAATTTGCAAACTCAGATGCTTCCAAAGCCATCAGTGGGAGACCGAGCACA 352
DB 6442 CTTTGTCAATTTGCAAACTCAGATGCTTCCAAAGCCATCAGTGGGAGACCGAGCACA 6501
QY 353 GGGAGGACCAAGGGAGAGAGAAAGAAATTAAGCAACGTTATTCTTAACACA 412
DB 6502 GGGAGGACCAAGGGAGAGAGAAAGAAATTAAGCAACGTTATTCTTAACACA 6560
QY 413 CTTTCTATAGAGAGTTGAGAGAGTGCATATTTTTTAAATCTCAGTGGCAATATTCA 472
DB 6561 CTTTCTATAGAGAGTTGAGAGAGTGCATATTTTTTAAATCTCAGTGGCAATATTCA 6620
QY 473 AAGTTTTCATGTGTCTTAACAAAGGTGT 501
DB 6621 AAGTTTTCATGTGTCTTAACAAAGGTGT 6649

RESULT 5
AL357553/c 155198 bp DNA linear HTG 10-JUL-2001
LOCUS AL357553
DEFINITION Homo sapiens chromosome 13 clone RP11-56D6, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
ACCESSION AL357553

VERSION AL357553.8 GI:13620383
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 155198)
AUTHORS Peck, A.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonequests@sanger.ac.uk
On Apr 12, 2001 this sequence version replaced gi:12956937.
COMMENT
Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: ba56D6
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 152143 bases at least Q40
Consensus quality: 153069 bases at least Q30
Consensus quality: 153660 bases at least Q20
Insert size: 155098; sum-of-ctrls
Insert size: 144134; 11.1% error; agarose-fp
Quality coverage: 5.59x in Q20 bases; sum-of-ctrls quality
coverage: 6.57x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2153: contig of 2153 bp in length
* 2154 2253: gap of 100 bp
* 2254 155198: contig of 152945 bp in length.
Location/Qualifiers
1. 155198
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-56D6"
/clone_1fb="RPCT-11.1"
1. 2153
/note="assembly_fragment:00588"
2254. 155198
/note="assembly_fragment:02529
clone_end:sp6
vector_side:right"

BASE COUNT 43663 a 34926 c 34669 g 41836 t 104 others
ORIGIN

Query Match 72.9%; Score 365; DB 2; Length 155198;
Best Local Similarity 100.0%; Pred. No. 6.2e-94;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 AGATCTGCCCCCTGGAGAGACAGAGGCTTTACCGAATTCCTTCAATCTTCAACG 196
DB 130593 AGATCTGCCCCCTGGAGAGACAGAGGCTTTACCGAATTCCTTCAATCTTCAACG 130534
QY 197 CCATCAGTGGAGACTCCAAACAGCAATGTTTACGAGATGACCAAGCTGTTGCTGCG 256
DB 130533 CCATCAGTGGAGACTCCAAACAGCAATGTTTACGAGATGACCAAGCTGTTGCTGCG 130474
QY 257 TGTGATTACATCTCAGTCCCGGTGTGTGGGACTTCTTGTCAATTTGCAAACTCAGAT 316
DB 130533 TGTGATTACATCTCAGTCCCGGTGTGTGGGACTTCTTGTCAATTTGCAAACTCAGAT 316

Db	Accession	Version	Source	Organism	Title	Journal	Comment
Db 130473	TGATATTAATCTCATGGCCCGTGTGTGGGACCTTGCTTTGTCATTTGCAAACTACGAT	130414					
QY	317	GCCTTCCAAAGCCAAATCACTGGGAGACCGACAGAGGAGGACCAAGGAGGAGAGA	376				
Db 130413	GCCTTCCAAAGCCAAATCACTGGGAGACCGACAGAGGAGGAGGAGGAGAGA	130354					
QY	377	GAAGAGAAATTAAGAACGCTTATTTCTTACAGACTTCTATAGGAGTTGTAAAGAG	436				
Db 130353	GAAGAGAAATTAAGAACGCTTATTTCTTACAGACTTCTATAGGAGTTGTAAAGAG	130294					
QY	437	TGCACATATTTTTTTTAAATCTCATCGCAATTTTCAAAAGTTTCATTGCTTAAACAA	496				
Db 130293	TGCACATATTTTTTTTAAATCTCATCGCAATTTTCAAAAGTTTCATTGCTTAAACAA	130234					
QY	497	GGTGT 501					
Db 130233	GGTGT 130229						
RESULT 6							
AL161420/c							
LOCUS	AL161420	163316 bp	DNA	linear	PRI 01-FEB-2001		
DEFINITION	Human DNA sequence from clone RP11-155N3 on chromosome 13 contains ESTs, STS and GSSs. Contains the 3' part of a novel gene similar to KIAA0694, the KIAA1058 gene and a putative novel gene, complete sequence.						
ACCESSION	AL161420						
VERSION	AL161420.10	GI:10443397					
KEYWORDS	HTG, KIAA0694, KIAA1058,						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.						
AUTHORS	1 (bases 1 to 163316)						
TITLE	Smith, M.						
JOURNAL	Direct Submission						
COMMENT	Submitted (31-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Requests: clonerequests@sanger.ac.uk On Oct 1, 2000 this sequence version replaced gi:10039669. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone configs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr13 This sequence is the entire insert of clone RP11-155N3 The true left end of clone RP11-318G11 is at 114983 in this sequence. The true right end of clone RP11-56b6 is at 42341 in this sequence. The true left end of clone RP11-551M18 is at 43464 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-155N3 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/Dacpac/home.htm VECTOR: pBACE3.6. Location/Qualifiers 1..163316 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="13"						

misc_feature	/clone="Rp11-155N3"
	/clone_11b="RpCt-11.1"
	6. .510
repeat_region	/note="match: GSS: Em:AQ372034"
	75. .372
repeat_region	/note="AluDb repeat: matches 1. .298 of consensus"
	447. .742
repeat_region	/note="AluSp repeat: matches 1. .296 of consensus"
	745. .1025
repeat_region	/note="L1MC1 repeat: matches 6019. .6327 of consensus"
	1250. .1566
repeat_region	/note="AluSk repeat: matches 2. .311 of consensus"
	1631. .2325
repeat_region	/note="L1MC3 repeat: matches 6683. .7400 of consensus"
	2326. .2619
repeat_region	/note="AluSk repeat: matches 1. .295 of consensus"
	2620. .2890
misc_feature	/note="L1MC5 repeat: matches 7400. .7665 of consensus"
	complement(3019. .3311)
misc_feature	/note="match: STS: Em:HSN273ZH1"
	3060. .3191
misc_feature	/note="match: GSS: Em:B17811"
	complement(3065. .3246)
misc_feature	/note="match: STS: Em:AF009212"
	3067. .3190
repeat_region	/note="match: GSS: Em:AQ114323"
	3087. .3122
repeat_region	/note="18 copies 2 mer tg 97% conserved"
	3089. .3124
repeat_region	/note="6 copies 6 mer tgggtg 97% conserved"
	3123. .3280
misc_feature	/note="match: GSS: Em:AQ459714"
	3191. .3475
repeat_region	/note="AluSk repeat: matches 1. .288 of consensus"
	3692. .3773
repeat_region	/note="MIR repeat: matches 3. .94 of consensus"
	3887. .4074
repeat_region	/note="L2 repeat: matches 1607. .1787 of consensus"
	4082. .4144
repeat_region	/note="3 copies 21 mer 88% conserved"
	4418. .4630
repeat_region	/note="L2 repeat: matches 1354. .1587 of consensus"
	7208. .7609
repeat_region	/note="MLT1B repeat: matches 8. .390 of consensus"
	9220. .9344
repeat_region	/note="MIR repeat: matches 71. .196 of consensus"
	9586. .9667
repeat_region	/note="L1M2 repeat: matches 5569. .5651 of consensus"
	9714. .9813
repeat_region	/note="L2 repeat: matches 771. .874 of consensus"
	9869. .10202
repeat_region	/note="L2 repeat: matches 971. .1261 of consensus"
	10203. .10581
repeat_region	/note="MST1 repeat: matches 1. .426 of consensus"
	10582. .11267
repeat_region	/note="L2 repeat: matches 1261. .2121 of consensus"
	11460. .11570
repeat_region	/note="L2 repeat: matches 2377. .2497 of consensus"
	11603. .11778
repeat_region	/note="L1MC3 repeat: matches 5996. .6166 of consensus"
	11779. .12160
repeat_region	/note="MER1A repeat: matches 1. .371 of consensus"
	12161. .12729
repeat_region	/note="L1MC3 repeat: matches 5403. .5996 of consensus"
	12682. .12950
repeat_region	/note="L1M6 repeat: matches 1. .274 of consensus"
	12945. .13209
repeat_region	/note="L1MC5 repeat: matches 7249. .7517 of consensus"
	13323. .13413
repeat_region	/note="AluU/FPM repeat: matches 221. .311 of consensus"
	13711. .13846
misc_feature	/note="AluSk/k repeat: matches 1. .136 of consensus"
	13955. .14417

Db	2007	AGTTTCATCCGTCAAGTTGTGTGACTTAACTTCATGAAACCAACAGCTGGGAGCAATGCTTTGTCA	2066
Qy	301	TTTTCGAAACTCAGAGATGCTTTCCAAAGCCATCTCATCTGGGAGACCAGAGCACAGAGAGGA-	359
Db	2067	TGTGCAAACTCAGAGAGACTTTCAGAGCTAAATCACTGTGTTTGTGCACAGCACAGGAAGAAG	2126
Qy	360	CCAAGGGGAAAGGGGACAGCAAAAGAAATAAAGAACACGTTATTTCTTAAACAGACTTTCTA	419
Db	2127	CCATGGGGAAATGGGAAAGGAAAGAGCCCTGAGACTGTGATTAATTAATGAGCA--TTTTA	2184
Qy	420	TAGAGATGTGAACAGAGTCACATATTTTTTTTAAATCTCACTGGCAATATTTCAAATTTT	479
Db	2185	TAGAGATGTGGGGAAGAAAGTGCACATATTTTTTTTAAATCTCACTGGCAATGTT--TAATTTT	2242
Qy	480	CATTGTGTCTTAAACAAAGGTG	501
Db	2243	CCCTCATGTCTTAAACAGGTGAT	2264

RESULT 9	AC126253	182640 bp	DNA	linear	HTG 04-JUL-2002
LOCUS	AC126253				WORKING DRAFT
DEFINITION	Mus musculus chromosome UKM clone RP24-119M15, JOURNAL				
ACCESSION	AC126253				
VERSION	AC126253.1	GI:21693957			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FUL1TOP.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 182640)				
TITLE	McPherson, J.D. and Waterston, R.H.				
JOURNAL	The sequence of Mus musculus clone unpublished				
REFERENCE	2 (bases 1 to 182640)				
AUTHORS	McPherson, J.D. and Waterston, R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (04-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
COMMENT					

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: W.BB019M15
----- Summary Statistics -----
Sequencing vector: M13: 0%
Sequencing vector: plasmid: 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 1778371 bases at least Q40
Consensus quality: 178021 bases at least Q30
Consensus quality: 178402 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 181640; sum-of-ontigs
Quality coverage: 11.19 in Q20 bases; agarose-fp
Quality coverage: 9.89 in Q20 bases; sum-of-ontigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4276: contig of 4276 bp in length

*	4376:	gap of unknown length
*	4377	8990: contig of 4614 bp in length
*	8991	9090: gap of unknown length
*	9091	14391: contig of 5301 bp in length
*	14392	14491: gap of unknown length
*	14492	23041: contig of 8550 bp in length
*	23042	23141: gap of unknown length
*	23142	37848: contig of 14707 bp in length
*	37849	37948: gap of unknown length
*	37949	52177: contig of 14229 bp in length
*	52177	52277: gap of unknown length
*	52278	75893: contig of 23616 bp in length
*	75893	75993: gap of unknown length
*	75993	113549: contig of 37556 bp in length
*	113549	113649: gap of unknown length
*	113650	177361: contig of 63712 bp in length
*	177362	177461: gap of unknown length
*	177462	179201: contig of 1740 bp in length
*	179202	179301: gap of unknown length
*	179302	182640: contig of 3339 bp in length

```

FEATURES
source
    location/Qualifiers
    1. .182640
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /chromosome="UNK"
    /clone="RP24-119M15"
    1. .4276
    /note="assembly_name:Contig10"
    4377. .8990
    /note="assembly_name:Contig11"
    9091. .14391
    /note="assembly_name:Contig12"
    14492. .23041
    /note="assembly_name:Contig13"
    23142. .37848
    /note="assembly_name:Contig14"
    37949. .52177
    /note="assembly_name:Contig15"
    52278. .75893
    /note="assembly_name:Contig16"
    75994. .113549
    /note="assembly_name:Contig17"
    113650. .177361
    /note="assembly_name:Contig18"
    177462. .179201
    /note="assembly_name:Contig18"
    179302. .182640
    /note="assembly_name:Contig19"
BASE COUNT    49101 a 42136 c 41309 g 49052 t 1042 other
ORIGIN

```

Query Match	44.0%	Score 220.6	DB 2	Length 182640
Best Local Similarity	75.4%	Pred. No. 2.8e-57		
Matches 301	Conservative	0	Mismatches 94	Indels 4
			Gaps 2	
QY 96	CAGGAAGTGGCAGAGAGCGCTTTCGAATCATGCAATGACGACATCTGCCCCCTGGAGGA	155		
Db 162621	CAGAGCAFTGCCCTTGATTAACATGTCTTCTGTACCTCGAGATTTGGCCCCCTGGAGGA	162680		
QY 156	GAACACAGCGCTCTTACCGAATTTCCCTTTACATCTTTCACAGCCCATCATGTTGGACTCCAAC	215		
Db 162681	GAACACAGCGTCTCTACCAATATCCCTGTGCACATCTTTCACAGCCCATCATGCGGGACACCAAC	162740		
QY 216	AAGCAATGTTTACGGGATGACAGCGTCGCTCGCTCGTGTGATTATCATCTCATGAC	275		
Db 162741	AAGCACATGTGTTCAAGGTTGACACAGCTGCTCTCAATGTGTGTGATTTTAACTCATGAA	162800		
QY 276	CCGTGTGTGGGAGCTGCTTTGTCTATTTTCAAACTCAGATGCTTTTCCAAAGCCATATCAC	335		
Db 162801	CCGTGTGTGGGAGATATGCTTTGTCTATGTCAAACTCAGATGATCTTCCAAAGCTATATCAC	162860		
QY 336	TGGGAGATCCGACACAGGAGGACCAAGGGAGAGGGAGAGAAAGCAATTAAGAACAA	395		

QY 456 CTCACCTGCATATTCGAAGTTTCATTGCTCTTACA 494
Db 84613 CTCACCTGCATATTCGAAGTTTCATTGCTCTTACA 84649

RESULT 11

AC109966

LOCUS

DEFINITION

AC109966 175281 bp DNA linear HTG 13-JUL-2002
Rattus norvegicus clone CH230-32315, *** SEQUENCING IN PROGRESS
*** 75 unordered pieces.

AC109966

AC109966.3 GI:21738210

HTG: HTGS.PHASEL

SOURCE

ORGANISM

Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 175281)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alstrooks,S.L., Amaralunga,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbata,J., Benson,J., Bimaga,K., Blankenburg,K., Bonin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Butlay,C., Burch,P., Burkelt,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Dayala,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinj,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.D.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homes,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,O., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,G., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Melker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokemkwo,S., Ogun,M., Okwunnu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherrer,S., Scott,G., Shen,H., Shoshari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tatney,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vilson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18847026.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: Q0P

Center clone name: CH230-32315

----- Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 105442 bases at least Q40

Consensus quality: 111492 bases at least Q30

Consensus quality: 116270 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a "working draft" sequence. It currently

* consists of 75 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1290: contig of 1290 bp in length
1291 1390: gap of unknown length
1391 2414: contig of 1024 bp in length
2414 2514: gap of unknown length
2514 2615: contig of 1125 bp in length
2615 3639: gap of unknown length
3639 3739: gap of unknown length
3739 5063: contig of 1324 bp in length
5063 5163: gap of unknown length
5163 5164: contig of 1080 bp in length
5164 6243: gap of unknown length
6243 6344: gap of 1193 bp in length
6344 7537: gap of unknown length
7537 7636: gap of unknown length
7636 9071: contig of 1435 bp in length
9071 9171: gap of unknown length
9171 10200: contig of 1029 bp in length
10200 10300: gap of unknown length
10300 11505: contig of 1205 bp in length
11505 11605: gap of unknown length
11605 12680: contig of 1075 bp in length
12680 12780: gap of unknown length
12780 14056: contig of 1276 bp in length
14056 14156: gap of unknown length
14156 15413: contig of 1257 bp in length
15413 15513: gap of unknown length
15513 15514: contig of 1067 bp in length
15514 16581: gap of unknown length
16581 17682: gap of 1012 bp in length
17682 17792: gap of unknown length
17792 19013: contig of 1221 bp in length
19013 19113: gap of unknown length
19113 20418: contig of 1305 bp in length
20418 20518: gap of unknown length
20518 21880: contig of 1362 bp in length
21880 21981: gap of unknown length
21981 22985: contig of 1005 bp in length
22985 23086: gap of unknown length
23086 24156: contig of 1071 bp in length
24156 24256: gap of unknown length
24256 25732: contig of 1476 bp in length
25732 25832: gap of unknown length
25832 27196: contig of 1364 bp in length
27196 27296: gap of unknown length
27296 28715: contig of 1319 bp in length
28715 30145: gap of unknown length
30145 30145: contig of 1430 bp in length


```
* 30146 30245: gap of unknown length
* 30246 31418: contig of 1173 bp in length
* 31419 31518: gap of unknown length
* 31519 32896: contig of 1378 bp in length
* 32897 32996: gap of unknown length
* 32997 34142: contig of 1146 bp in length
* 34143 34242: gap of unknown length
* 34243 35262: contig of 1020 bp in length
* 35263 35924: gap of unknown length
* 35925 37024: contig of 1562 bp in length
* 37025 38670: gap of unknown length
* 38671 38770: contig of 1646 bp in length
* 38771 40612: gap of unknown length
* 40613 40713: contig of 1842 bp in length
* 40713 42724: contig of 2012 bp in length
* 42725 42825: gap of unknown length
* 42825 43923: contig of 1005 bp in length
* 43923 43930: gap of unknown length
* 43930 45819: contig of 1889 bp in length
* 45819 48183: gap of unknown length
* 48183 48283: contig of 2265 bp in length
* 48283 48887: gap of unknown length
* 48887 49888: contig of 1604 bp in length
* 49888 49988: gap of unknown length
* 49988 51668: contig of 1581 bp in length
* 51668 53251: gap of unknown length
* 53251 53351: contig of 1583 bp in length
* 53351 55636: gap of unknown length
* 55636 55737: contig of 2285 bp in length
* 55737 57159: gap of unknown length
* 57159 57259: contig of 1423 bp in length
* 57259 59519: gap of unknown length
* 59519 59620: contig of 2260 bp in length
* 59620 63393: gap of unknown length
* 63393 63493: contig of 3773 bp in length
* 63493 65130: gap of unknown length
* 65130 65230: contig of 1638 bp in length
* 65230 68145: gap of unknown length
* 68145 68245: contig of 2915 bp in length
* 68245 70071: gap of unknown length
* 70071 70170: contig of 1825 bp in length
* 70170 73494: gap of unknown length
* 73494 73594: contig of 3324 bp in length
* 73594 75991: gap of unknown length
* 75991 76091: contig of 2397 bp in length
* 76091 77867: gap of unknown length
* 77867 77968: contig of 1776 bp in length
* 77968 80319: gap of unknown length
* 80319 80419: contig of 2351 bp in length
* 80419 82211: gap of unknown length
* 82211 82311: contig of 1793 bp in length
* 82311 84512: gap of unknown length
* 84512 84513: contig of 2101 bp in length
* 84513 85982: gap of unknown length
* 85982 86083: contig of 1470 bp in length
* 86083 86083: gap of unknown length
* 86083 88448: gap of 2266 bp in length
* 88448 90746: gap of unknown length
* 90746 90847: contig of 2298 bp in length
* 90847 92600: gap of unknown length
* 92600: contig of 1754 bp in length
```

Query Match 41.28; Score 206.2; DB 2; Length 175281;
Best Local Similarity 74.78; Pired. No. 4e-48;
Matches 286; Conservative 0; Mismatches 93; Indels 4; Gaps 2;

```
QY 119 CTGAATCATCTTCACAGATGTCGCCCTGGAGAGAGAGAGCGCTTACCGAATT 178
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133646 CTGCTTCTTGCTACCTCAAGATTGGCCCTGGAGAGAGAGAAAGGCTCTCAAAATT 133705
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 CCTTCACATCTTCACAGCGCATGATGGAGCTCCAAACACACAAATGTTACGGGATGA 238
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 133706 CCTTCACATCTTCACAGCGCATGATGGAGCGACCAACACAGATGTTCAAGGTTGA 133765
QY 239 CCAGCTGCTTGGGTGCTGATTTACATCTCATGAGCCCGTGTGGGACCTTGTGT 298
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133766 CCAGTTCATCCTCAGTTGTGTGACTTACTGACCAACACAGCTGGGACATGCTTGT 133825
QY 299 CATTTGCAACTCAGGATGTTTCCAAAGCCATCCTCTGGGAGACCGACAGGAGAG 358
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133826 CATGTGAAACTCAGAGCGACTTTCAGAGCTAAATCCTGATTTTGGCAAGCAGGAAGA 133885
QY 359 ACCAAGGAGAGGGGAGAGAGAAAGAAATTAAGCAACGTTATTTCTTAACAGACTTGT 418
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133886 ACCCATGGGAATGGGAGAGAGAAAGAGCCCTGGACCTGTGATTTAATAGCAGA--TTTT 133943
QY 419 ATAGGACTTTGAAGAGGTGCACATATTTTTTAAATTCACCTGGCAATTTCAAGTTT 478
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133944 ATAGGACTTTGGGAGAGGTGCACATATTTTTTAAATTCACCTGGCAAGTTT--TAGTTT 134001
QY 479 TCATTTGCTCTTAACAAAGGTGT 501
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 134002 TCCTCATGCTTTAACAGGTGTAT 134024
```

```
RESULT 12
AC119357/c
LOCUS
DEFINITION
Rattus norvegicus clone CH250-473M19, *** SEQUENCING IN PROGRESS
*** 89 unordered pieces.
ACCESSION
AC119357
VERSION
AC119357.3 GI:21746916
KEYWORDS
HTG; HTGS; PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 192825)
REFERENCE
1 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshan,F.R., Allen,C.,
Alsprinks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,F.,
Barbarta,J., Benton,J., Blincoe,K., Blumentburg,K., Bonini,D.,
Bouck,U., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Burch,P., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Datborne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Hornel,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karissom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisied,H.,
Lozado,R.D., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mashiney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenko,S., Oguni,M., Okunodu,G.,
Orangunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swatek,A., Tabors,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umanli,K., Vasquez,L., Vera,V., Villalton,D., Vinson,R., Wang,O.,
```

TITLE
JOURNAL
Unpublished
2 (bases 1 to 192825)
Worley, K. C.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (26-Apr-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 192825)
Worley, K. C.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (18-Jul-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20429771.
COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GVEP
Center clone name: CH230-473M19

Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 117932 bases at least Q40
Consensus quality: 124308 bases at least Q30
Consensus quality: 129203 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 89 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1130 1129: contig of 1129 bp in length
* 1230 1229: gap of unknown length
* 2850 2849: contig of 1620 bp in length
* 2950 2949: gap of unknown length
* 4126 4125: contig of 1176 bp in length
* 4226 4225: gap of unknown length
* 5645 5644: contig of 1419 bp in length
* 5745 5744: gap of unknown length
* 6998 6997: contig of 1253 bp in length
* 7098 7097: gap of unknown length
* 8758 8757: contig of 1660 bp in length
* 8858 8857: gap of unknown length
* 9861 9860: contig of 1003 bp in length
* 11606 11605: contig of 1645 bp in length
* 11706 11705: gap of unknown length
* 12752 12751: contig of 1046 bp in length
* 12852 12851: gap of unknown length
* 14491 14490: contig of 1639 bp in length
* 14591 14590: gap of unknown length
* 16102 16101: contig of 1511 bp in length
* 16202 16201: gap of unknown length
* 17629 17628: contig of 1427 bp in length
* 17729 17728: gap of unknown length
* 18883 18883: contig of 1155 bp in length
* 18984 20211: gap of unknown length
* 20212 20311: gap of unknown length
20312 21353: contig of 1042 bp in length
21354 21453: gap of unknown length
21454 22518: contig of 1065 bp in length
22519 22618: gap of unknown length
22619 24107: contig of 1489 bp in length
24108 24208: gap of unknown length
24208 25603: contig of 1396 bp in length
25604 25703: gap of unknown length
25704 27171: contig of 1468 bp in length
27172 27272: gap of unknown length
27272 28993: contig of 1722 bp in length
28993 29093: gap of unknown length
29094 30122: contig of 1029 bp in length
30123 30222: gap of unknown length
30223 31317: contig of 1095 bp in length
31318 31417: gap of unknown length
31418 33144: contig of 1721 bp in length
33145 33244: gap of unknown length
33245 34431: contig of 1187 bp in length
34432 34531: gap of unknown length
34532 35743: contig of 1212 bp in length
35744 35843: gap of unknown length
35844 37180: contig of 1337 bp in length
37181 37280: gap of unknown length
37281 38596: contig of 1316 bp in length
38597 38696: gap of unknown length
38697 40031: contig of 1335 bp in length
40032 40131: gap of unknown length
40132 41485: contig of 1354 bp in length
41486 41585: gap of unknown length
41586 42744: contig of 1159 bp in length
42745 42844: gap of unknown length
42845 45158: contig of 2314 bp in length
45159 45258: gap of unknown length
45259 46752: contig of 1494 bp in length
46753 46852: gap of unknown length
46853 48029: contig of 1177 bp in length
48029 48129: gap of unknown length
48130 49682: contig of 1553 bp in length
49683 49782: gap of unknown length
49783 51724: contig of 1942 bp in length
51725 51824: gap of unknown length
51825 53137: contig of 1313 bp in length
53138 53237: gap of unknown length
53238 55287: contig of 2050 bp in length
55288 55387: gap of unknown length
55388 57026: contig of 1639 bp in length
57027 57126: gap of unknown length
57127 58750: contig of 1624 bp in length
58751 60180: gap of unknown length
60180 60180: contig of 1330 bp in length
60181 60280: gap of unknown length
60281 61971: contig of 1691 bp in length
61972 62071: gap of unknown length
62072 63893: contig of 1822 bp in length
63894 63993: gap of unknown length
63994 65782: contig of 1789 bp in length
65783 65882: gap of unknown length
65883 67125: contig of 1243 bp in length
67126 69485: contig of 2260 bp in length
69486 71428: contig of 1843 bp in length
71429 71528: gap of unknown length
71529 73230: contig of 1702 bp in length
73231 73330: gap of unknown length
73331 75374: contig of 2044 bp in length
75375 75474: gap of unknown length
75475 77927: contig of 2453 bp in length
77928 78027: gap of unknown length
78028 79268: contig of 1241 bp in length
79269 81963: gap of unknown length
81963: contig of 2595 bp in length

Takao Isogai, Helix Research Institute,

International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

Location/Qualifiers
1. .2036

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000139"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_idb="NT2RM4"
/note="cloning vector: pME18SFJ3-mRNA from uninduced NT2
neuronal precursor cells."
153..1400

```

```

/note="unnamed protein product"
/codon_start=1
/protseq_id="BAA91583.1"
/db_xref="GI:7022394"
/translation="MFSMGPAFLSTIPNLIKEGAMKEDSGMDPTYNENITVEQLYWM
CVELPEKSRREYLADVCPPIIAVEKORDRKRKSLDYLDIHRSLVKAENVSEKRL
FRVRYRAVYGGGFEDEEGKKYIKERKLLSLISQRLKLVADFGANVY.I.TODD
SKRVKPDLDPEYAAITQYLYVYFPEEKELIDRKDFEMHINIRNVEFETLGGKK
HSGVAEOCKRPIALITQYLYVYFPEKKRIVIOISDRKDEIVAIIDEMSKKVELNOLC
TMEEDVMDRLQIKLOGSYGVKNAGMPMYAARFLEFNARKYYPNQYKLLKEITRQRA
DAGGALDYNNELIKEDQLEVOEELRSTKDKMLSTELVYNNQELITGRDLSKRVDQDT
CTRYLSKAPALPELVYSSISAAV"

```

Best Local Similarity 67.5%; Pred. No. 2,1e-11;
Matches 110; Conservative 0; Mismatches 53; Indels 0

Matches	110;	Conservative	0;	Mismatches	53;	Indels	0;	Caps	0;
---------	------	--------------	----	------------	-----	--------	----	------	----

1 CAATTGAGGAGCTTGGGGTCAAGCTTATAGCGGTAACCGAAGCTTCGATTAAAGAAGAC 60

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

[illegible]

BC015018

	2719 bp	mRNA	linear	FBI 04 OCT - 2001
DEFINITION	Homo sapiens, clone MGC:8871 IMAGE:3922744 , mRNA, complete cds.			

VERSION BC015018.1 GI:15929124

SOURCE Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Placentalia; Carnivora; Canidae; Felidae; Felis; Felis concolor

REFERENCE
AUTHORS
1 (bases 1 to 2413)
Strausberg, R.

JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammary

ТАБЛИЦА ТРАНСЛИТАЦИИ / ЗНАЧЕНИЯ СЛОВА	ТАБЛИЦА / ЗНАЧЕНИЯ СЛОВА	ТАБЛИЦА / ЗНАЧЕНИЯ СЛОВА
---------------------------------------	--------------------------	--------------------------

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
contact: MGC help desk

CDNA Library Prepared by: The I.M.A.G.E. Consortium (LBNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Raelle Jones, Benjamin Asano, Ian Boscoe, Aaron Butterfield, Jetticia Hsiao, Martin Krzywinski, Beta Kutsche, Oliver Li

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy,
Ness, Pawan Pandoh, Anna-Liisa Prabhui, parvaneh Saeedi,

Schein, Duane Smailus, Michael Smith, Lorraine Spence, Je-
 Michael Thorne, Miranada Tsai, Natassja van den Bosch, Jill
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 15 Row: p Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8923209.

Location/Qualifiers
1..2413

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:8871 IMAGE:392274"
/tissue="type="skin, melanotic melanoma."
/clone_lib="NIR_MGC_72"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
139.1167

```

```

/codon_start=1
/product="Unknown (protein for MGC:8871)"
/protein_id="AAH15018.1"
/db_xref="GI:15929125"
/translation="MNSNPNPAEVRKDLTRKIPVLNATAQOMEHEKDEPMVLQVSYST
ANSVSTPELRRTVETLSMAKIHARNQDLSAANCYIHIALILEYLRKRGYKVEKIC
TASLSEDTHPCDNSNLTTPSGSFMFSCWMPALSTTPINKEGKAGEGSDGDDTPK
NNINIVEQLYMCVEFLMEKSERELIADVNPILAVEKEGKALISDLYVDIHSYGLK
VAEVSNEERLGRYRYVAFYGGGFEDEGKEIKYPEKLTGSETSORLKLTYADK
RADWNIILQDSNKNVNRKDLDEPVATIQWTYVYPRFEKIEIDRKTFEEMHNINRRPY
FETPPTLSGKKGGAEOCKRRITLTSHLEPYVKRIQVISOSTETLNPLEVAIDEX
SKVSELDNLCEMEVDMLRLOLKLQGSVKNVNAAPMAARFLEETPNKKYDQNVQNG
KLKEETRFQADACGALDVNERLIKEDOLEYOEELSHYKQDMLSELSTVWNEQITGH
DQLSRGVQGTCTRYTSKATPAPATYSISSAEV"

```

Query Match	DB 9	Length	2413
121	Score 78.2	DB 9	Length 2413
163	Score 78.2	DB 9	Length 2413

ATGCGTCTGGAGTTACGGGCAAGCAGCCGTCGA	1308
Matches	110; Conservative 0; Mismatches 53; Indels 0; Gaps 0

1 CAAATTGAGGAACCTTCCGGTCAAGCCCTTAGCCGTAACCGAACGTCGATTAAAGAAACAC 60

BC015018	2413 bp	mRNA	linear	PRI	04-OCT-2001
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9
10	10	10	10	10	10
11	11	11	11	11	11
12	12	12	12	12	12
13	13	13	13	13	13
14	14	14	14	14	14
15	15	15	15	15	15
16	16	16	16	16	16
17	17	17	17	17	17
18	18	18	18	18	18
19	19	19	19	19	19
20	20	20	20	20	20
21	21	21	21	21	21
22	22	22	22	22	22
23	23	23	23	23	23
24	24	24	24	24	24
25	25	25	25	25	25
26	26	26	26	26	26
27	27	27	27	27	27
28	28	28	28	28	28
29	29	29	29	29	29
30	30	30	30	30	30
31	31	31	31	31	31
32	32	32	32	32	32
33	33	33	33	33	33
34	34	34	34	34	34
35	35	35	35	35	35
36	36	36	36	36	36
37	37	37	37	37	37
38	38	38	38	38	38
39	39	39	39	39	39
40	40	40	40	40	40
41	41	41	41	41	41
42	42	42	42	42	42
43	43	43	43	43	43
44	44	44	44	44	44
45	45	45	45	45	45
46	46	46	46	46	46
47	47	47	47	47	47
48	48	48	48	48	48
49	49	49	49	49	49
50	50	50	50	50	50
51	51	51	51	51	51
52	52	52	52	52	52
53	53	53	53	53	53
54	54	54	54	54	54
55	55	55	55	55	55
56	56	56	56	56	56
57	57	57	57	57	57
58	58	58	58	58	58
59	59	59	59	59	59
60	60	60	60	60	60
61	61	61	61	61	61
62	62	62	62	62	62
63	63	63	63	63	63
64	64	64	64	64	64
65	65	65	65	65	65
66	66	66	66	66	66
67	67	67	67	67	67
68	68	68	68	68	68
69	69	69	69	69	69
70	70	70	70	70	70
71	71	71	71	71	71
72	72	72	72	72	72
73	73	73	73	73	73
74	74	74	74	74	74

[illegible]

KEYWORDS: *carcinoma, MG, MGCC, MGCCF, MGCCF*

[illegible]

REFERENCE
1 (bases 1 to 2413)
Mammalia; Eumetazoa; Platyhelminthes; Cnidaria; Mollusca; Arthropoda; Insecta; Hymenoptera; Homoptera; Hominiidae; Hominoidea; Homo.

TITLE	
Direct Submision	

08-06-27 N.	
search completed:	February 11, 2003, 08:11:28
Job time :	1465 66 secs

Gene Collection (MGC), Cancer Genomics Office, National Cancer

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1: SEQ ID NO 6595; 1399pp + Sequence listing: English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoietic regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 1493 BP; 466 A; 314 C; 249 G; 464 T; 0 other;

Query Match 100.0%; Score 501; DB 22; Length 1493;
Best Local Similarity 100.0%; Pred. No. 2.1e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTTGTGGAAGCTTGGGTCAAGCCTTACGGTAAACGAACTGTGATTAAGAGAC 60
DB 1404 CAATTTGTGGAAGCTTGGGTCAAGCCTTACGGTAAACGAACTGTGATTAAGAGAC 1345
QY 61 CAGCTCGAGTATCAGGAAGAAATGAAGCCAACTACAGGAAATGGCGAGACCTTTCT 120
DB 1344 CAGCTCGAGTATCAGGAAGAAATGAAGCCAACTACAGGAAATGGCGAGACCTTTCT 1285
QY 121 GAAATCAGCATGACGACATCTGCCCTGGAGAGAGACGAGGCTTCCGCAATTC 180
DB 1284 GAAATCAGCATGACGACATCTGCCCTGGAGAGAGAGACGAGGCTTCCGCAATTC 1225
QY 181 CTTCACATCTTCAGGCGCATAGTGGGAGCTCAACAGACACAGATGGTTACAGGGATGACC 240
DB 1224 CTTCACATCTTCAGGCGCATAGTGGGAGCTCAACAGACACAGATGGTTACAGGGATGACC 1165
QY 241 ACCTGCTCTGGGCTGTGATTAATCTCATGCCCCGTGTGTGGGACCTGCTTTGCA 300
DB 1164 ACCTGCTCTGGGCTGTGATTAATCTCATGCCCCGTGTGTGGGACCTGCTTTGCA 1105
QY 301 TTTCGAATCTCGAGATGCTTTCCAAAGCCAAATCACTGGGAGACGAGACAGGAGAC 360
DB 1104 TTTCGAATCTCGAGATGCTTTCCAAAGCCAAATCACTGGGAGACGAGACAGGAGAC 1045
QY 361 CAAGGGAGAGGGAGAGAAAGAAATAAAGAACAGCTTATTTCTTAACAGACTTCTAT 420
DB 1044 CAAGGGAGAGGGAGAGAAAGAAATAAAGAACAGCTTATTTCTTAACAGACTTCTAT 985
QY 421 AGGAGTTGTAAGAGGTGCACATATTTTAAATCTCACTGGGCAATTTCAAAATTTTC 480
DB 984 AGGAGTTGTAAGAGGTGCACATATTTTAAATCTCACTGGGCAATTTCAAAATTTTC 925
QY 481 ATTGTCTTAAACAAGGTGT 501
DB 924 ATTGTCTTAAACAAGGTGT 904

RESULT 2
ABK84967
ID ABK84967 standard: cDNA; 3614 BP.
XX
AC ABK84967;
XX
DT 13-AUG-2002 (first entry)
XX

DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #2.
XX
XX Human; autoimmune disease; haematopoietic disorder; Digorge syndrome;
KW blood protein disorder; agammaglobulinemia; dysgammaglobulinemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-0532202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
PI WPI; 2002-416861/44.
DR P-PSDB; AB616173.
XX
XX New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure: Figure 3a; 245pp; English.
XX
XX The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of Th1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinemia,
CC dysgammaglobulinemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or

CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.

XX Sequence 3614 BP; 1009 A; 834 C; 874 G; 897 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 3614;
Best Local Similarity 100.0%; Pred. No. 3.3e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTTGGGAAGCTTGGGCTCAAGCTTACCGTAAACGACCTGATTAAGAAGAC 60
DB 2607 CAATTTGGGAAGCTTGGGCTCAAGCTTACCGTAAACGACCTGATTAAGAAGAC 2666
QY 61 CAGCTGAGATTCAGGAAGAAATGAAGCAACTACAGGAATGGCGAAGAGCTTCT 120
DB 2667 CAGCTGAGATTCAGGAAGAAATGAAGCAACTACAGGAATGGCGAAGAGCTTCT 2726
QY 121 GAATCATGATGAGCAGATCTGCCCTCGAGAGAAAGACGACGCTTACCGAATTC 180
DB 2727 GAATCATGATGAGCAGATCTGCCCTCGAGAGAAAGACGACGCTTACCGAATTC 2786
QY 181 CTTCACATCTTCAAGCCATCAGTGGGACCTCCACACAGCAATGCTTACAGGGATGAC 240
DB 2787 CTTCACATCTTCAAGCCATCAGTGGGACCTCCACACAGCAATGCTTACAGGGATGAC 2846
QY 241 AGCTGCTCTTGGCTGCTGTGATTACATCTCATGCGCCGCTGTGGGACCTTGTGCA 300
DB 2847 AGCTGCTCTTGGCTGCTGTGATTACATCTCATGCGCCGCTGTGGGACCTTGTGCA 2906
QY 301 TTTCGAAATCAGATGCTTCCAAAGCAATCACTGGGAGACCGACAGAGGAGGAC 360
DB 2907 TTTCGAAATCAGATGCTTCCAAAGCAATCACTGGGAGACCGACAGAGGAGGAC 2966
QY 361 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 2967 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3026
QY 421 AGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 3027 AGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3086
QY 481 ATTGTGCTTAAACAAAGGTGT 501
DB 3087 ATTGTGCTTAAACAAAGGTGT 3107

RESULT 3
ABK84964
ID ABK84964 standard; cDNA; 4806 BP.
XX
AC ABK84964;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP).
XX
KW Human; autoimmune disease; haematopoietic disorder; Digesorge syndrome;
KW blood protein disorder; agammaglobulinemia; dysgamma globulinemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anapylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.

PF 15-OCT-2001; 2001WO-US32202.
XX
XX 13-OCT-2000; 2000US-0687837.
PR
XX
XX (ARBO-) ARBOR VITA CORP.
PA
XX (GARW-) GARMAN J D.
PA
XX (CAND-) CANDIA A F.

XX
XX Lu PS;
XX
XX MPI: 2002-416861/44.
DR
XX P-PSDB; ABG61670.
DR
XX

XX New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
XX an immune response, and for treating multiple sclerosis, rheumatoid
XX arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
XX and sepsis -
XX

PS Disclosure: Figure 1; 245pp; English.

XX The invention relates to an isolated polypeptide (I) comprising an amino
XX acid sequence that has 90 % sequence identity to one of the human
XX cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
XX sequences (PS). (I) is useful for identifying a compound or agent that
XX binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
XX detecting a CLASP-2 polypeptide in a sample. (III) is useful for
XX inhibiting an immune response in a subject. A pharmaceutical composition
XX comprising a nucleic acid encoding (I), or (II) is useful for preventing
XX or treating a CLASP-2 mediated disease e.g., an autoimmune disease, where
XX the autoimmune disease is caused or exacerbated by increased activity
XX of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
XX detection or inhibition of CLASP-2 expression (e.g., antisense or
XX ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
XX polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
XX antibodies or are used as therapeutic polypeptides. The CLASP-2
XX polynucleotide or fragments can be used in diagnostics (e.g., as probes
XX for CLASP-2 expression), as a lymphocyte marker and for therapeutic
XX purposes. CLASP-2 polynucleotides can construct transgenic and knockout
XX animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
XX polynucleotides can screen for CLASP-2 agonists and antagonists.
XX CLASP-2 polypeptides or polynucleotides can treat deficiencies or
XX disorders of the immune system, by activating or inhibiting the
XX activation, differentiation of immune cells and can treat or detect
XX deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
XX or polynucleotides can increase differentiation and proliferation of
XX haematopoietic cells, including the pluripotent stem cells to treat those
XX disorders associated with a decrease in certain (or many) types of
XX haematopoietic cells e.g., immunologic deficiency syndromes including
XX blood protein disorders (e.g., agammaglobulinemia,
XX dysgamma globulinemia, ataxia telangiectasia, common variable
XX immunodeficiency, Digesorge syndrome, lymphopenia, thrombocytopenia, or
XX haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
XX detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
XX Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
XX endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
XX inflammation. CLASP-2 can be used to treat anapylaxis or
XX hypersensitivity to an antigenic molecules, organ rejection or graft-
XX versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
XX represent cadherin-like asymmetry protein (CLASP) coding sequences and
XX PCR primers of the invention.

SO Sequence 4806 BP; 1352 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4806;
Best Local Similarity 100.0%; Pred. No. 3.8e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTTGGGAAGCTTGGGCTCAAGCTTACCGTAAACGACCTGATTAAGAAGAC 60
DB 3799 CAATTTGGGAAGCTTGGGCTCAAGCTTACCGTAAACGACCTGATTAAGAAGAC 3858
QY 61 CAGCTGAGATTCAGGAAGAAATGAAGCAACTACAGGAATGGCGAAGAGCTTCT 120
|||||

Db 3859 CAGCTCGAGTATCAGGAAGAAATGAAGCCAACTACAGGGAATGGCAGAGACTTTCT 3918
QY 121 GAAATCATGATGAGCAGATCTGCCCCCTGGAGAGAAAGACGCTTTACGAATTC 180
Db 3919 GAAATCATGATGAGCAGATCTGCCCCCTGGAGAGAAAGACGCTTTACGAATTC 3978
QY 181 CTTGCATCTTCAAGCCATCAGTGGGACTCCACACACAAATGGTTCACGAGATGCC 240
Db 3979 CTTGCATCTTCAAGCCATCAGTGGGACTCCACACACAAATGGTTCACGAGATGCC 4038
QY 241 ACCTGCTCTGGCTGCTGATTTACATCTCATGCGCCCTGTGTGGGACTGCTTTGTCA 300
Db 4039 ACCTGCTCTGGCTGCTGATTTACATCTCATGCGCCCTGTGTGGGACTGCTTTGTCA 4098
QY 301 TTTCGAAACTCAGATGCTTTTCCAAAGCCAACTACTGSGGAGACCGACACAGGAGGAC 360
Db 4099 TTTCGAAACTCAGATGCTTTTCCAAAGCCAACTACTGSGGAGACCGACACAGGAGGAC 4158
QY 361 CAAGGGGAAGGGAGAGAAAGAAATGAAGCAACGTTATTTCTTACAGACTTCTAT 420
Db 4159 CAAGGGGAAGGGAGAGAAAGAAATGAAGCAACGTTATTTCTTACAGACTTCTAT 4218
QY 421 AGGAGTTGTAAGAGTGCACATATTTTAAATCTCAGTGGCAATATTCAAAGTTTC 480
Db 4219 AGGAGTTGTAAGAGTGCACATATTTTAAATCTCAGTGGCAATATTCAAAGTTTC 4278
QY 481 ATTGTGCTTTAACAAGGTGT 501
Db 4279 ATTGTGCTTTAACAAGGTGT 4299

RESULT 4

AC87972
ID AAC87972 standard; cDNA; 4807 BP.

AC87972;

07-MAR-2001 (first entry)

Human CLASP-2 nucleotide sequence.

XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
XX cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
XX immunomodulatory; antiinflammatory; antiarthritic; cytoskeletal;
XX hypotensive; antithrombotic; antihypertensive; haemostatic; neuroprotective;
XX hypersensitivity; transplantation rejection response; immunodeficiency;
XX proliferation; differentiation; inflammatory response; arthritis;
XX inflammatory bowel disease; haematopoietic cell; blood protein disorder;
XX anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
XX endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
PN WO200061747-A2.
PD 19-OCT-2000.
PE 13-APR-2000; 2000MO-US10158.
PF
XX
XX 14-APR-1999; 9905-0129171.
XX 14-MAY-1999; 9905-0134114.
XX 14-MAY-1999; 9905-0134117.
XX 14-MAY-1999; 9905-0134118.
XX 21-OCT-1999; 9905-0160860.
XX 29-OCT-1999; 9905-0162498.
XX 13-DEC-1999; 9905-0170453.
XX 14-JAN-2000; 2000US-0176195.
XX 14-FEB-2000; 2000US-0182296.

(ARBO-) ARBOR VITA CORP.

Lu PS:

XX
XX
XX
XX
XX

DR WPI: 2000-619230/59.
DR P-PSDB; AAB36527.

XX Isolated cadherin-1-like asymmetry protein-2 polynucleotide and
XX polypeptide used to diagnose, treat and prevent autoimmune diseases and
XX inflammatory responses -
XX

Example 1; Fig 1; 286pp; English.

XX The present invention describes cadherin-1-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antiarthritic, cytoskeletal, hypotensive, antithrombotic,
XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX immune response in a subject may also be inhibited by administering an
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX proteins and antibodies can be used to prevent or treat a CLASP-2
XX mediated disease, such as an autoimmune disease caused or exacerbated
XX by increased activity of TH1 cells. They can also be used to treat
XX hypersensitivities, prevent transplantation rejection responses and
XX augment immune responsiveness in immunodeficiency states, inhibit
XX proliferation and differentiation of cells involved in an inflammatory
XX response e.g., arthritis, inflammatory bowel disease and increase
XX differentiation and proliferation of haematopoietic cells e.g. to treat
XX anaemia, thrombocytopenia and other blood protein disorders. Disorders
XX treated by disrupting CLASP-2 function include multiple sclerosis,
XX rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
XX The present sequence encodes human CLASP-2, which is used in the
XX exemplification of the present invention.

SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4807;

Best Local Similarity 100.0%; Pred No. 3.8e-143; Mismatches 0; Gaps 0;

Matches 501; Conservative 0; Indels 0;

QY 1 CAATTTGTGGAAGCTTCGCGTCAAGCCTTAGCGGTAAAGCAAGCTGTGATTAAGAAGAC 60
Db 3800 CAATTTGTGGAAGCTTCGCGTCAAGCCTTAGCGGTAAAGCAAGCTGTGATTAAGAAGAC 3859
QY 61 CAGCTCGAGTATCAGGAAGAAATGAAGCCAACTACACAGGGAATGGGAGGAGCTTCT 120
Db 3860 CAGCTCGAGTATCAGGAAGAAATGAAGCCAACTACACAGGGAATGGGAGGAGCTTCT 3919
QY 121 GAAATCATGATGAGCAGATCTGCCCCCTGGAGAGAAAGACGCTTTACGAATTC 180
Db 3920 GAAATCATGATGAGCAGATCTGCCCCCTGGAGAGAAAGACGCTTTACGAATTC 3979
QY 181 CTTGCATCTTCAAGCCATCAGTGGGACTCCACACACAAATGGTTCACGAGATGCC 240
Db 3980 CTTGCATCTTCAAGCCATCAGTGGGACTCCACACACAAATGGTTCACGAGATGCC 4039
QY 241 AGCTGCTCTGGCTGCTGATTTACATCTCATGCGCCCTGTGTGGGACTGCTTTGTCA 300
Db 4040 AGCTGCTCTGGCTGCTGATTTACATCTCATGCGCCCTGTGTGGGACTGCTTTGTCA 4099
QY 301 TTTCGAAACTCAGATGCTTTTCCAAAGCCAACTACTGSGGAGACCGACACAGGAGGAC 360
Db 4100 TTTCGAAACTCAGATGCTTTTCCAAAGCCAACTACTGSGGAGACCGACACAGGAGGAC 4159
QY 361 CAAGGGGAAGGGAGAGAAAGAAATGAAGCAACGTTATTTCTTACAGACTTCTAT 420
Db 4160 CAAGGGGAAGGGAGAGAAAGAAATGAAGCAACGTTATTTCTTACAGACTTCTAT 4219
QY 421 AGGAGTTGTAAGAGTGCACATATTTTAAATCTCAGTGGCAATATTCAAAGTTTC 480
Db 4220 AGGAGTTGTAAGAGTGCACATATTTTAAATCTCAGTGGCAATATTCAAAGTTTC 4279
QY 481 ATTGTGCTTTAACAAGGTGT 501
Db 4280 ATTGTGCTTTAACAAGGTGT 4300

RESULT 5
AAC87973
ID AAC87973 standard; cDNA; 4807 BP.
XX
AC AAC87973;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CIASP-2A nucleotide sequence.
XX
KW CIASP-1; CIASP-2; transmembrane protein; immune response; inflammatory;
caderlin-like asymmetric protein; autoimmune disease; immunosuppressive;
immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometritis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN MO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
DR P-PDB: AAB36528.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Example 1; Fig 2B; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CIASP-2). CIASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CIASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CIASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CIASP-2. CIASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CIASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states. Inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g. arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CIASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometritis and pregnancy induced hypertension.
CC The present sequence encodes human CIASP-2A, which is used in the
CC exemplification of the present invention.

XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other:
XX
Query Match 100.0%; Score 501; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3,8e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAATTTGTGAGAGCTTGGGTCACGCTTACGCGTAAGCAAGCTGTGATTAAGAAGAC 60
|||||
Db 3800 CATTTTGTGGAAGCTTGGCGTCAACGCTTACGCGTAAGCAAGCTGTGATTAAGAAGAC 3859
XX
QY 61 CAGCTCGAGTATCAGAGAAATGAAGCCAACTACAGGAAATGGCGAAGAGCTTTCT 120
|||||
Db 3860 CAGCTCGAGTATCAGAGAAATGAAGCCAACTACAGGAAATGGCGAAGAGAGCTTTCT 3919
XX
QY 121 GAATCATGCATGAGAGATGTCGCCCTGGAGGAGAGAGAGAGAGAGCTTACGAATTC 180
|||||
Db 3920 GAATCATGCATGAGAGATGTCGCCCTGGAGGAGAGAGAGAGAGAGCTTACGAATTC 3979
XX
QY 181 CTTCACATCTTCACGCCCATCAGTGGACTCCACACAGCACAATGTTTCACGGAGTACC 240
|||||
Db 3980 CTTCACATCTTCACGCCCATCAGTGGACTCCACACAGCACAATGTTTCACGGAGTACC 4039
XX
QY 241 AGCTGCTTGGTGCTGTGATTCATCATGAGCCGTGTGGGAGCTTTGTGCA 300
|||||
Db 4040 AGCTGCTTGGTGCTGTGATTCATCATGAGCCGTGTGGGAGCTTTGTGCA 4099
XX
QY 301 TTTGCAACTCAGATGCTTTCCAAAGCCAACTACGTGGGAGACCGAGCAGAGGAGAC 360
|||||
Db 4100 TTTGCAACTCAGATGCTTTCCAAAGCCAACTACGTGGGAGACCGAGCAGAGGAGAC 4159
XX
QY 361 CAAGGGAGAGGGAGAGAAATGAAGCAATTAATCTTCAACAGCTTCTAT 420
|||||
Db 4160 CAAGGGAGAGGGAGAGAAATGAAGCAATTAATCTTCAACAGCTTCTAT 4219
XX
QY 421 AGGAGTTGTAGAGAGTGCACATATTTTAAATCTCAGTGGCAATTTCAAGTTTTC 480
|||||
Db 4220 AGGAGTTGTAGAGAGTGCACATATTTTAAATCTCAGTGGCAATTTCAAGTTTTC 4279
XX
QY 481 ATTGTGCTTAACAAAGTGT 501
|||||
Db 4280 ATTGTGCTTAACAAAGTGT 4300
XX
RESULT 6
ABK84966
ID ABK84966 standard; cDNA; 4807 BP.
XX
AC ABK84966;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CIASP) isoform #1.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgamma globulinemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopaenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometritis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
CIASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX

CC acid sequence that has 90 % sequence identity to one of the human
 CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2a, 2b, 2c, 2e)
 CC sequences (PS). (I) is useful for identifying a compound or agent that
 CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
 CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
 CC inhibiting a immune response in a subject. A pharmaceutical composition
 CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
 CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
 CC the autoimmune disease is caused or exacerbated by increased activity
 CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
 CC detection or inhibition of CLASP-2 expression (e.g., antisense or
 CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
 CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
 CC antibodies or are used as therapeutic polypeptides. The CLASP-2
 CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
 CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
 CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
 CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
 CC polynucleotides can screen for CLASP-2 agonists and antagonists.
 CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
 CC disorders of the immune system, by activating or inhibiting the
 CC activation, differentiation of immune cells and can treat or detect
 CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
 CC or polynucleotides can increase differentiation and proliferation of
 CC haematopoietic cells, including the pluripotent stem cells to treat those
 CC disorders associated with a decrease in certain (or many) types of
 CC haematopoietic cells e.g., immunologic deficiency syndromes including
 CC blood protein disorders (e.g., agammaglobulinemia,
 CC dysgammaaglobulinemia, ataxia telangiectasia, common variable
 CC immunodeficiency, DiGeorge syndrome, lymphopenia, thrombocytopenia, or
 CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
 CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
 CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
 CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
 CC inflammation. CLASP-2 can be used to treat anaphylaxis or
 CC hypersensitivity to an antigenic molecules, organ rejection or graft-
 CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
 CC represent cadherin-like asymmetry protein (CLASP) coding sequences
 CC and PCR primers of the invention.

XX
 SO Sequence 4807 BP: 1353 A: 1117 C: 1136 G: 1201 T: 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4807;
 Best Local Similarity 100.0%; Pred. No. 3.8e-143;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTGTGGAGCTGGCGTCAAGCCTTAGGGTAACGACGTCGATTAAGAAGAC 60
 DB 3800 CAATTGTGGAGCTGGCGTCAAGCCTTAGGGTAACGACGTCGATTAAGAAGAC 3859
 QY 61 CAGCTCGAGTATCAGGAGAATGAAGCCACTACAGGAAATGGCGAGAGCTTCT 120
 DB 3860 CAGCTCGAGTATCAGGAGAATGAAGCCACTACAGGAAATGGCGAGAGCTTCT 3919
 QY 121 GAATATCATGATGACGATCTGCCCTCGAGAGAGAGAGAGCGTCTTACCGAATCC 180
 DB 3920 GAATATCATGATGACGATCTGCCCTCGAGAGAGAGAGAGCGTCTTACCGAATCC 3979
 QY 181 CTTTCACATCTTCAACGCCATTCAGTGGACTCCACAAGACACATGTTTCAGGAGTAC 240
 DB 3980 CTTTCACATCTTCAACGCCATTCAGTGGACTCCACAAGACACATGTTTCAGGAGTAC 4039
 QY 241 AGCTCGTCTTGGCTGCTGATTTACATCTCATGCGCCGCTGTGGGACTGCTTGTCA 300
 DB 4040 AGCTCGTCTTGGCTGCTGATTTACATCTCATGCGCCGCTGTGGGACTGCTTGTCA 4099
 QY 301 TTGGCAAACTCAGATGCTTTCCAAAGCCATACCTGGGAGAGACGACACAGAGAGAC 360
 DB 4100 TTGGCAAACTCAGATGCTTTCCAAAGCCATACCTGGGAGAGACGACACAGAGAGAC 4159
 QY 361 CAAAGGGAAGGGAAGAAAGAAATGAAGCAACGTTATTTCTTAACAGACTTCTAT 420
 DB 4160 CAAAGGGAAGGGAAGAAAGAAATGAAGCAACGTTATTTCTTAACAGACTTCTAT 4219

QY 421 AGAGTTGTGAAGAGAGTGCACATATTTTTTAATCTCACTGGCAATATCAAGTTTC 480
 DB 4220 AGAGTTGTGAAGAGAGTGCACATATTTTTTAATCTCACTGGCAATATCAAGTTTC 4279
 QY 481 ATTGTGTCTTAACAAGGTGT 501
 DB 4280 ATTGTGTCTTAACAAGGTGT 4300

RESULT 8

MAC87974
 ID AAC87974 standard; cDNA: 4898 BP.

MAC87974;
 AC AAC87974;

DT 07-MAR-2001 (first entry)

XX Preliminary CLASP-2 nucleotide sequence #1.

XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
 XX cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
 XX immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
 XX hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
 XX hypersensitivity; transplantation rejection response; immunodeficiency;
 XX proliferation; differentiation; inflammatory response; arthritis;
 XX inflammatory bowel disease; haematopoietic cell; blood protein disorder;
 XX anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
 XX endometriosis; pregnancy induced hypertension; ss.

XX Homo sapiens.

XX MO200061747-A2.

XX 19-OCT-2000.

XX 13-APR-2000; 2000WO-US10158.

XX 14-APR-1999; 99US-0129171.

XX 14-MAY-1999; 99US-0134114.

XX 14-MAY-1999; 99US-0134117.

XX 14-MAY-1999; 99US-0134118.

XX 21-OCT-1999; 99US-0160860.

XX 29-OCT-1999; 99US-0162498.

XX 13-DEC-1999; 99US-0170453.

XX 14-JAN-2000; 2000US-0176195.

XX 14-FEB-2000; 2000US-0182296.

XX (ARBO-) ARBOR VITA CORP.

XX Lu PS;

XX WPI: 2000-619230/59.

XX P-PSDB: AAB36529.

XX Isolated cadherin-like asymmetry protein-2 polynucleotide and

XX polypeptide used to diagnose, treat and prevent autoimmune diseases and

XX inflammatory responses -

XX disclosure; Fig 10A; 286pp; English.

CC The present invention describes cadherin-like asymmetry protein-2
 CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
 CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
 CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
 CC used to inhibit an immune response in a subject by interfering with the
 CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
 CC immune response in a subject may also be inhibited by administering an
 CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
 CC proteins and antibodies can be used to prevent or treat a CLASP-2
 CC mediated disease, such as an autoimmune disease caused or exacerbated
 CC by increased activity of TH1 cells. They can also be used to treat
 CC hypersensitivities, prevent transplantation rejection responses and

```
CC      augment immune responsiveness in immunodeficiency states, inhibit
CC      proliferation and differentiation of cells involved in an inflammatory
CC      response e.g, arthritis, inflammatory bowel disease and increase
CC      differentiation and proliferation of haematopoietic cells e.g. to treat
CC      anaemia, thrombocytopenia and other blood protein disorders. Disorders
CC      treated by disrupting CLASP-2 function include multiple sclerosis,
CC      rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC      The present sequence represents a preliminary CLASP-2 nucleotide
CC      sequence, from the present invention.
CC      XX
SQ      Sequence 4898 BP: 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match      100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3,9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAATTGTGGAAGCTGGGGTCAAGCCTTAGCGGTAAGAAAGCAAGCTGATTAAAGAGAC 60
DB      3891 CAATTGTGGAAGCTTGGCGTCAAGCCTTAGCGGTAAGAAAGCAAGCTGATTAAAGAGAC 3950
QY      61 CAGCTCGAGTATCAGGAAGAATGAAGCCAACTACAGGGAAATGGCGAAGAGCTTTCT 120
DB      3951 CAGCTCGAGTATCAGGAAGAATGAAGCCAACTACAGGGAAATGGCGAAGAGCTTTCT 4010
QY      121 GAAATCATGATGAGCAGATCTGCCCCCTGGAGGAGAAAGACAGCGCTGTTACCGAATTCC 180
DB      4011 GAAATCATGATGAGCAGATCTGCCCCCTGGAGGAGAAAGACAGCGCTGTTACCGAATTCC 4070
QY      181 CTTGACATCTTCAAGCCCATCAAGTGGAGCTCCAAACAGACAAATGTTACCGGATGACC 240
DB      4071 CTTGACATCTTCAAGCCCATCAAGTGGAGCTCCAAACAGACAAATGTTACCGGATGACC 4130
QY      241 AGCTGCTCTGGTGTGTGATTACATCTCATGCCCCGTGTGTGGGAGCTTGTGTCA 300
DB      4131 AGCTGCTCTGGTGTGTGATTACATCTCATGCCCCGTGTGTGGGAGCTTGTGTCA 4190
QY      301 TTGGCAACTCAGATGCTTTCGCAAGCCCAATCACTGGGGAGACGAGACAGGGAGAGC 360
DB      4191 TTGGCAACTCAGATGCTTTCGCAAGCCCAATCACTGGGGAGACGAGACAGGGAGAGC 4250
QY      361 CAAGGGGAAGGGAGAGAAAGAAATGAAGCAACAGTTATTTCTTAACAGACTTTCTAT 420
DB      4251 CAAGGGGAAGGGAGAGAAAGAAATGAAGCAACAGTTATTTCTTAACAGACTTTCTAT 4310
QY      421 AGGAGTTTGAAGAGTGCACATATTTTAAATCTCCTGCGCAATATTCAAAGTTTC 480
DB      4311 AGGAGTTTGAAGAGTGCACATATTTTAAATCTCCTGCGCAATATTCAAAGTTTC 4370
QY      481 ATGTGCTTAAACAAGGTGT 501
DB      4371 ATGTGCTTAAACAAGGTGT 4391

RESULT 9
AAC87975
ID      AAC87975 standard; cDNA; 4898 BP.
AC      AAC87975;
XX
XX
XX      07-MAR-2001 (first entry)
DE      Preliminary CLASP-2 nucleotide sequence #2.
XX
XX      CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
XX      cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
XX      immunomodulatory; antiinflammatory; antiarthritic; cytoskeletal;
XX      hypotensive; antirheumatic; antianaemic; haemostatic; neutroprotective;
XX      hypersensitivity; transplantation rejection response; immunodeficiency;
XX      proliferation; differentiation; inflammatory response; arthritis;
XX      inflammatory bowel disease; haematopoietic cell; blood protein disorder;
XX      anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
XX      endometriosis; pregnancy induced hypertension; ss.

OS      Homo sapiens.
XX
XX      WO200061747-A2.
XX
XX      19-OCT-2000.
XX
XX      13-APR-2000; 2000WO-US10158.
XX
XX      14-APR-1999; 99US-0129171.
XX      14-MAY-1999; 99US-0134114.
XX      14-MAY-1999; 99US-0134117.
XX      14-MAY-1999; 99US-0134118.
XX      21-OCT-1999; 99US-0160860.
XX      29-OCT-1999; 99US-0162498.
XX      13-DEC-1999; 99US-0170453.
XX      14-JAN-2000; 2000US-0176195.
XX      14-FEB-2000; 2000US-0182296.
XX
XX      (ARBO-) ARBOR VITA CORP.
XX
XX      Lu PS:
XX
XX      WPI: 2000-619230/59.
XX
XX      Isolated cadherin-like asymmetry protein-2 polynucleotide and
XX      polypeptide used to diagnose, treat and prevent autoimmune diseases and
XX      inflammatory responses -
XX
XX      Disclosure; Fig 10B; 286pp; English.
XX
XX      The present invention describes cadherin-like asymmetry protein-2
XX      (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX      antiinflammatory, antiarthritic, cytoskeletal, hypotensive, antirheumatic,
XX      antianaemic, haemostatic and neutroprotective activities. CLASP-2 can be
XX      used to inhibit an immune response in a subject by interfering with the
XX      ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX      immune response in a subject may also be inhibited by administering an
XX      antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX      proteins and antibodies can be used to prevent or treat a CLASP-2
XX      mediated disease, such as an autoimmune disease caused or exacerbated
XX      by increased activity of TH1 cells. They can also be used to treat
XX      hypersensitivities, prevent transplantation rejection responses and
XX      augment immune responsiveness in immunodeficiency states, inhibit
XX      proliferation and differentiation of cells involved in an inflammatory
XX      response e.g, arthritis, inflammatory bowel disease and increase
XX      differentiation and proliferation of haematopoietic cells e.g. to treat
XX      anaemia, thrombocytopenia and other blood protein disorders. Disorders
XX      treated by disrupting CLASP-2 function include multiple sclerosis,
XX      rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
XX      The present sequence represents a preliminary CLASP-2 nucleotide
XX      sequence, from the present invention.
XX
XX      Sequence 4898 BP: 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match      100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3,9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAATTGTGGAAGCTGGGGTCAAGCCTTAGCGGTAAGAAAGCAAGCTGATTAAAGAGAC 60
DB      3891 CAATTGTGGAAGCTTGGCGTCAAGCCTTAGCGGTAAGAAAGCAAGCTGATTAAAGAGAC 3950
QY      61 CAGCTCGAGTATCAGGAAGAATGAAGCCAACTACAGGGAAATGGCGAAGAGCTTTCT 120
DB      3951 CAGCTCGAGTATCAGGAAGAATGAAGCCAACTACAGGGAAATGGCGAAGAGCTTTCT 4010
QY      121 GAAATCATGATGAGCAGATCTGCCCCCTGGAGGAGAAAGACAGCGCTTACCGAATTCC 180
DB      4011 GAAATCATGATGAGCAGATCTGCCCCCTGGAGGAGAAAGACAGCGCTTACCGAATTCC 4070
QY      181 CTTGACATCTTCAAGCCCATCAAGTGGAGCTCCAAACAGACAAATGTTACCGGATGACC 240
DB      4071 CTTGACATCTTCAAGCCCATCAAGTGGAGCTCCAAACAGACAAATGTTACCGGATGACC 4130
```


XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cyostatic;
KW hypotensive; antineumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
PN WO200061747-A2.
PD 19-OCT-2000.
XX
PE 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS:
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10D; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cyostatic, hypotensive, antineumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states. Inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g. arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTGTGGAAGCTTCGCGTCAAGCCTTAGCGGTAACGAACGCTGATTAAAGAAC 60
DB 3891 CAATTGTGGAAGCTTCGCGTCAAGCCTTAGCGGTAACGAACGCTGATTAAAGAAC 3950

QY 61 CAGCTCAGATATCAGAGGAAGAAATGAAGCCAACTACAGAGGAATGGCGAAGACCTTTCT 120
DB 3951 CAGCTCAGATATCAGAGGAAGAAATGAAGCCAACTACAGAGGAATGGCGAAGACCTTTCT 4010
QY 121 GAAATCATGATGAGCAGATCTGCCCTTGAGAGAGAACGAGCGTCTTACGCAATTC 180
DB 4011 GAAATCATGATGAGCAGATCTGCCCTTGAGAGAGAACGAGCGTCTTACGCAATTC 4070
QY 181 CTTCACATCTTCACAGCCCATCAGTGGGACCTCCACAGCAACAAATGTTCCAGGATGACC 240
DB 4071 CTTCACATCTTCACAGCCCATCAGTGGGACCTCCACAGCAACAAATGTTCCAGGATGACC 4130
QY 241 AGCTGCTCTTGCGTCTGTGATTAATCTCATGCGCCCTGTGTGGGACCTTGTCTCA 300
DB 4131 AGCTGCTCTTGCGTCTGTGATTAATCTCATGCGCCCTGTGTGGGACCTTGTCTCA 4190
QY 301 TTTGCAAACTCAGATGCTTTTCCAAAGCCAACTACCTGGGAGACCGACAGGAGGAC 360
DB 4191 TTTGCAAACTCAGATGCTTTTCCAAAGCCAACTACCTGGGAGACCGACAGGAGGAC 4250
QY 361 CAAGGGGAAGGGGAGAGAAAGAAATAAGAACACGTTATTCTTAAACAGACTTCTAT 420
DB 4251 CAAGGGGAAGGGGAGAGAAAGAAATAAGAACACGTTATTCTTAAACAGACTTCTAT 4310
QY 421 AGGAGTTGTAGAGGTCACATATTTTAAATCTCACCTGCGCAATATTCAAAGTTTC 480
DB 4311 AGGAGTTGTAGAGGTCACATATTTTAAATCTCACCTGCGCAATATTCAAAGTTTC 4370
QY 481 ATTGTGCTTAACAAAAGGTGT 501
DB 4371 ATTGTGCTTAACAAAAGGTGT 4391

RESULT 12

AAC87978
ID AAC87978 standard; cDNA: 4898 BP.AAC87978;
AC AAC87978;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #5.
XX

CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cyostatic;
KW hypotensive; antineumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
KW
XX
OS Homo sapiens.
PN WO200061747-A2.
PD 19-OCT-2000.
XX
PE 13-APR-2000; 2000WO-US10158.
XX

PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX

XX Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;
SQ
Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAATTTGTGGAAGCTTGGCTGCAAGCCTTAGCGGTAAACGACGCTGTGATTAAAGAAGAC 60
DB 3891 CAATTTGTGGAAGCTTGGCTGCAAGCCTTAGCGGTAAACGACGCTGTGATTAAAGAAGAC 3950
QY 61 CAGCTCGAGTATCAAGAGAATAAAGCCAACTACAGGGAATGGGAGGAGCTTTCT 120
DB 3951 CAGCTCGAGTATCAAGAGAATAAAGCCAACTACAGGGAATGGGAGGAGCTTTCT 4010
QY 121 GAAATCATGATGAGAGATGTCGCCCTTGAGAGAGAAGACGAGCTGTTACCGAATTC 180
DB 4011 GAAATCATGATGAGAGATGTCGCCCTTGAGAGAGAAGACGAGCTGTTACCGAATTC 4070
QY 181 CTTCACATCTTCAAGCCATCAGTGGGACCTCCAAACAGCACATGTTTCACGGATGACC 240
DB 4071 CTTCACATCTTCAAGCCATCAGTGGGACCTCCAAACAGCACATGTTTCACGGATGACC 4130
QY 241 AGCTGCTTGGCTGCTGCTGATTCATGTCATGGCCGCTGTGGGAGCTTGTGCA 300
DB 4131 AGCTGCTTGGCTGCTGCTGATTCATGTCATGGCCGCTGTGGGAGCTTGTGCA 4190
QY 301 TTTGCAAACTCAGATGCTTTCAAAGCCAACTCAGTGGGAGACGAGCACAGGAGAGAC 360
DB 4191 TTTGCAAACTCAGATGCTTTCAAAGCCAACTCAGTGGGAGACGAGCACAGGAGAGAC 4250
QY 361 CAGGCGGAAGGGGAGAGAAAGAAATAAGAACACGTTATTCTTAACAGACTTTCTAT 420
DB 4251 CAGGCGGAAGGGGAGAGAAAGAAATAAGAACACGTTATTCTTAACAGACTTTCTAT 4310
QY 421 AGGAGTTGTAAGAGTCCACATATTTTAAATCTCAGTGGGAATATTCAGAGTTTTC 480
DB 4311 AGGAGTTGTAAGAGTCCACATATTTTAAATCTCAGTGGGAATATTCAGAGTTTTC 4370
QY 481 ATTGTGCTTAAACAAGGTGT 501
DB 4371 ATTGTGCTTAAACAAGGTGT 4391

RESULT 14
AAC87980
ID AAC87980 standard; cDNA; 4898 BP.
XX
XX AAC87980;
AC
XX
XX
DT 07-MAR-2001 (first entry)
XX
XX Preliminary CLASP-2 nucleotide sequence #7.
DE
XX
XX CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
XX
XX cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
XX
XX immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
XX
XX hypotensive; antirheumatic; antianemic; haemostatic; neuroprotective;
XX
XX hypersensitivity; transplantation rejection response; immunodeficiency;
XX
XX proliferation; differentiation; inflammatory response; arthritis;
XX
XX inflammatory bowel disease; haematopoietic cell; blood protein disorder;
XX
XX anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
XX
XX endometriosis; pregnancy induced hypertension; ss.
OS
XX Homo sapiens.
XX
XX WO200061747-A2.
PN
XX
XX 19-OCT-2000.
PD
XX
XX 13-APR-2000; 2000MO-US10158.
PF
XX
XX 14-APR-1999; 9905-0129171.

PR 14-MAY-1999; 9905-0134114.
PR 14-MAY-1999; 9905-0134117.
PR 14-MAY-1999; 9905-0134118.
PR 21-OCT-1999; 9905-0160860.
PR 29-OCT-1999; 9905-0162498.
PR 13-DEC-1999; 9905-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
XX (ARBO-) ARBOR VITA CORP.
PA
XX
XX Lu PS;
PI
XX WPI; 2000-619230/59.
DR
XX
XX Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
XX
XX Disclosure; Fig 10G; 286pp; English.
XX
XX The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH1 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states. Inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
XX
SQ Sequence 4898 BP; 1379 A; 1134 C; 1166 G; 1219 T; 0 other;
Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAATTTGTGGAAGCTTGGCTGCAAGCCTTAGCGGTAAACGACGCTGTGATTAAAGAAGAC 60
DB 3891 CAATTTGTGGAAGCTTGGCTGCAAGCCTTAGCGGTAAACGACGCTGTGATTAAAGAAGAC 3950
QY 61 CAGCTCGAGTATCAAGAGAATAAAGCCAACTACAGGGAATGGGAGGAGCTTTCT 120
DB 3951 CAGCTCGAGTATCAAGAGAATAAAGCCAACTACAGGGAATGGGAGGAGCTTTCT 4010
QY 121 GAAATCATGATGAGAGATGTCGCCCTTGAGAGAGAAGACGAGCTGTTACCGAATTC 180
DB 4011 GAAATCATGATGAGAGATGTCGCCCTTGAGAGAGAAGACGAGCTGTTACCGAATTC 4070
QY 181 CTTCACATCTTCAAGCCATCAGTGGGACCTCCAAACAGCACATGTTTCACGGATGACC 240
DB 4071 CTTCACATCTTCAAGCCATCAGTGGGACCTCCAAACAGCACATGTTTCACGGATGACC 4130
QY 241 AGCTGCTTGGCTGCTGCTGATTCATGTCATGGCCGCTGTGGGAGCTTGTGCA 300
DB 4131 AGCTGCTTGGCTGCTGCTGATTCATGTCATGGCCGCTGTGGGAGCTTGTGCA 4190
QY 301 TTTGCAAACTCAGATGCTTTCAAAGCCAACTCAGTGGGAGACGAGCACAGGAGAGAC 360
DB 4191 TTTGCAAACTCAGATGCTTTCAAAGCCAACTCAGTGGGAGACGAGCACAGGAGAGAC 4250

QY 361 CAAGGGGAAGGGAGAGAAAGAAATTAAGAACACGTTATTTCTTAACAGACTTTCTAT 420
CC |
CC |
CC |
DB 4251 CAAGGGGAAGGGAGAGAAAGAAATTAAGAACACGTTATTTCTTAACAGACTTTCTAT 4310
QY 421 AGAGAGTTGAAGAGGTGCACATATTTTAACTCTACGCGCAATTTCAAGCTTTTC 480
CC |
CC |
CC |
DB 4311 AGAGAGTTGAAGAGGTGCACATATTTTAACTCTACGCGCAATTTCAAGCTTTTC 4370
QY 481 ATTGTGCTTTAACAAAGGTGT 501
CC |
CC |
DB 4371 ATTGTGCTTTAACAAAGGTGT 4391
RESULT 15
AAC87981
ID AAC87981 standard; cDNA; 4898 BP.
XX
AC AAC87981;
XX
XX 07-MAR-2001 (first entry)
DE Preliminary CLASP-2 nucleotide sequence #8.
XX
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antirheumatic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
XX Homo sapiens.
OS
XX
XX MO200061747-A2.
FN
PD 19-OCT-2000.
XX
XX 13-APR-2000; 2000MO-US10158.
FE
XX 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX
XX Lu PS;
XX
XX WPI: 2000-619230/59.
DR
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
XX
PS Disclosure; Fig 10H; 286pp; English.
XX
XX
XX The present invention describes cadherin-like asymmetry protein-2
XX (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
XX antiinflammatory, antirheumatic, cytostatic, hypotensive, antirheumatic,
XX antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
XX used to inhibit an immune response in a subject by interfering with the
XX ability of a CLASP-2 protein to bind to another T cell or B cell. An
XX immune response in a subject may also be inhibited by administering an
XX antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
XX proteins and antibodies can be used to prevent or treat a CLASP-2
XX mediated disease, such as an autoimmune disease caused or exacerbated

CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e.g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other:
Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAATTGTGAGAGCTTGGCGTCAAGCCTTAGCGGTTAAGCAAGCTGCTGAATTAAGAGAC 60
DB 3891 CAATTGTGAGAGCTTGGCGTCAAGCCTTAGCGGTTAAGCAAGCTGCTGAATTAAGAGAC 3950
QY 61 CAGCTCAGTATCAGGAAGAATGAAGCCAACTACAGGGAATGCGAAGAGCTTTCT 120
DB 3951 CAGCTCAGTATCAGGAAGAATGAAGCCAACTACAGGGAATGCGAAGAGCTTTCT 4010
QY 121 GAATATCATGATAGACAGATCTCCCTTGGAGAGAGACGAGCTTTACCGAATTC 180
DB 4011 GAATATCATGATAGAGAGATCTCCCTTGGAGAGAGACGAGCTTTACCGAATTC 4070
QY 181 CTTCACATCTTCAAGCCATCACTGGAGCTTCAACAGCAATGGTTACGGGATGACC 240
DB 4071 CTTCACATCTTCAAGCCATCACTGGAGCTTCAACAGCAATGGTTACGGGATGACC 4130
QY 241 AGCTCGTCTTGGTGTGATTAATCATGATGCCGTGTGTGGGACTTCTTGCA 300
DB 4131 AGCTCGTCTTGGTGTGATTAATCATGATGCCGTGTGTGGGACTTCTTGCA 4190
QY 301 TTTCGAACCTCAGGATGCTTTTCCAAAGCCATCATCTGGGGAACGAGCAGGAGAC 360
DB 4191 TTTCGAACCTCAGGATGCTTTTCCAAAGCCATCATCTGGGGAACGAGCAGGAGAC 4250
QY 361 CAAGGGGAAGGGAGAGAAAGAAATTAAGAACACGTTATTTCTTAACAGACTTTCTAT 420
DB 4251 CAAGGGGAAGGGAGAGAAAGAAATTAAGAACACGTTATTTCTTAACAGACTTTCTAT 4310
QY 421 AGAGAGTTGAAGAGGTGCACATATTTTAACTCTACGCGCAATTTCAAGCTTTTC 480
DB 4311 AGAGAGTTGAAGAGGTGCACATATTTTAACTCTACGCGCAATTTCAAGCTTTTC 4370
QY 481 ATTGTGCTTTAACAAAGGTGT 501
DB 4371 ATTGTGCTTTAACAAAGGTGT 4391

Search completed: February 7, 2003, 07:08:42
Job time : 163.085 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:00:34 : Search time 28.6858 Seconds
(without alignments)
5356.145 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300

Perfect score: 501 caattgtggaagcttcgag.....tgtgtcttaacaagtggt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 862724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	27.7	1031	4 US-09-397-787-157	Sequence 157, Appl
2	41.8	8.3	2325	4 US-09-484-970B-33	Sequence 33, Appl
3	40.8	8.1	7218	1 US-08-232-463-14	Sequence 14, Appl
4	33	6.6	36519	3 US-08-923-137-2	Sequence 2, Appl
5	32.8	6.5	2570	2 US-09-056-075-2	Sequence 2, Appl
6	32.8	6.5	2613	4 US-09-255-829-7	Sequence 7, Appl
7	32.8	6.5	3891	1 US-08-480-604A-27	Sequence 27, Appl
8	32.8	6.5	3891	4 US-08-405-496A-27	Sequence 27, Appl
9	32.8	6.5	3891	4 US-08-915-136-27	Sequence 27, Appl
10	32.4	6.5	289	4 US-09-007-005-17	Sequence 17, Appl
11	32.4	6.5	289	4 US-09-244-796-17	Sequence 17, Appl
12	32.4	6.5	6559	4 US-09-234-186-1	Sequence 1, Appl
13	32.4	6.5	6559	4 PCF-US93-05651-1	Sequence 1, Appl
14	32.4	6.5	6560	5 PCF-US93-05651-1	Sequence 39, Appl
15	32	6.4	1241	1 US-08-471-033-39	Sequence 42, Appl
16	32	6.4	1241	1 US-08-471-033-42	Sequence 39, Appl
17	32	6.4	1241	2 US-08-471-044-35	Sequence 35, Appl
18	32	6.4	1241	2 US-08-471-044-42	Sequence 42, Appl
19	32	6.4	1241	2 US-08-463-483A-39	Sequence 39, Appl
20	32	6.4	1241	2 US-08-463-483A-42	Sequence 42, Appl
21	32	6.4	1241	2 US-08-471-046A-39	Sequence 39, Appl
22	32	6.4	1241	2 US-08-471-046A-42	Sequence 42, Appl
23	32	6.4	1241	2 US-08-470-566B-39	Sequence 42, Appl
24	32	6.4	1241	2 US-08-470-566B-42	Sequence 42, Appl
25	32	6.4	1241	2 US-08-469-334-39	Sequence 39, Appl
26	32	6.4	1241	2 US-08-469-334-42	Sequence 42, Appl
27	32	6.4	1241	3 US-09-300-529-39	Sequence 39, Appl

28	32	6.4	1241	3 US-09-300-529-42	Sequence 42, Appl
29	32	6.4	1358	1 US-08-471-033-45	Sequence 45, Appl
30	32	6.4	1358	2 US-08-471-044-45	Sequence 45, Appl
31	32	6.4	1358	2 US-08-463-483A-45	Sequence 45, Appl
32	32	6.4	1358	2 US-08-471-046A-45	Sequence 45, Appl
33	32	6.4	1358	2 US-08-470-566B-45	Sequence 45, Appl
34	32	6.4	1358	2 US-08-469-334-45	Sequence 45, Appl
35	32	6.4	1358	1 US-09-300-529-45	Sequence 45, Appl
36	32	6.4	1389	1 US-08-471-033-27	Sequence 27, Appl
37	32	6.4	1389	2 US-08-471-044-27	Sequence 27, Appl
38	32	6.4	1389	2 US-08-463-483A-27	Sequence 27, Appl
39	32	6.4	1389	2 US-08-471-046A-27	Sequence 27, Appl
40	32	6.4	1389	2 US-08-470-566B-27	Sequence 27, Appl
41	32	6.4	1389	2 US-08-469-334-27	Sequence 27, Appl
42	32	6.4	1389	2 US-09-300-529-27	Sequence 27, Appl
43	32	6.4	1399	1 US-08-471-033-24	Sequence 24, Appl
44	32	6.4	1399	2 US-08-471-044-24	Sequence 24, Appl
45	32	6.4	1399	2 US-08-463-483A-24	Sequence 24, Appl

ALIGNMENTS

```
RESULT 1
US-09-397-787-157
; Sequence 157, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 157
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-157

Query Match      27.7% Score 139; DB 4; Length 1031;
Best Local Similarity 96.6%; Pred. No. 4.3e-33;
Matches 142; Conservative 0; Mismatches 5; Indels 0;

OY 1 CAATTGTGGAAGCTTGGCGTCAAGCCTTAAAGCAAGCTGTGATTAAAGAAC 60
    |||||||
Db 511 CAATTGTGGAAGCTTGGCGTCAAGCCTTAAAGCAAGCTGTGATTAAAGAAC 570

OY 61 CACCTGAGATCAGGAAGAAATGAAGCACTACAGGAATGCGAAGACCTTCT 120
    |||||||
Db 571 CACCTGAGATCAGGAAGAAATGAAGCACTACAGGAATGCGAAGACCTTCT 630

OY 121 GAATCATGATGAGCAGATCTGCC 147
    |||||||
Db 631 GAATCATGATGAGCAGATCTGCC 657

RESULT 2
US-09-484-970B-33
; Sequence 33, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkunth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
```

```

: CURRENT FILING DATE: 2000-01-18
: NUMBER OF SEQ ID NOS: 172
: SOFTWARE: PERL Program
: SEQ ID NO 33
: LENGTH: 2225
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. 6426186 238544.2CB1
US-09-484-970B-33

Query Match
Best Local Similarity 56.0%; Pred. No. 0.0035;
Matches 79: Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 16 TCCGTCAGACCTTAGCGGTAAACCAAGCTGTGATTAAAGAACCCAGCTCGAGATCAG 75
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 721 TGTGTGAAGCTAGAGAAAACAAGCGTCTCATCAGCGCAACCAAGGGAATATCAG 780

QY 76 GAAGAATGAAGCCACTACGAGGAATGGCGAAGAGCTTTCTGAATCATGCATGAG 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 781 CAGGAACCTAATAAGACATATACAGCTAATAAGAACTCAGGCCAATGATCGAGCGG 840

QY 136 CAGATCTGCCCTCGAGAGAG 156
   || || || || || || || || || || || || || || || || || || || ||
DB 841 AAAATTCCAGAACTGTACAG 861

RESULT 3
US-08-232-463-14/C
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/2,114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
```

```

: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZgpt-F15
: US-08-232-463-14

Query Match
Best Local Similarity 6.8%; Pred. No. 0.012;
Matches 30: Conservative 213; Mismatches 195; Indels 0; Gaps 0;

QY 8 TGGAGCTTCGCGTCAAGCCTTAGCGTAAACGACGCTGTATTAAAGACCAAGCTCG 67
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1450 TAGAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1391

QY 68 AGATATCAGAGAAATGAAGCCACTACAGGAATGGCGAAGAGCTTTCTGAATCA 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1390 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1331

QY 128 TGCATGACGATCTGCCCTGAGAGAGACGAGCGCTTACCGAATTCCTTGACA 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1330 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1271

QY 188 TCTTCAAGCCATCAGTGAGCTCAACAAGCAATGTTTACGGATGACGCTCGT 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1270 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1211

QY 248 CTTCGTCGTGTATTATCATCTCATGCGCCGCTGTGGGAGCTTGTGTATTTGCA 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1210 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1151

QY 308 ACTCAGATGCTTCCAAAGCCATCACTGGGAGAGACCGAGACAGGAGCAAGGG 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1150 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1091

QY 368 AAGGGAGAGAAAGAAATTAAGACAGCTTATTCTTACAGACTTCTATAGAGTT 427
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1090 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1031

QY 428 GTAGAAGTGCACATAT 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1030 TTAATTCTGTGAGCGTAT 1013

RESULT 4
US-08-923-137-2/C
: Sequence 2, Application US/08923137
: Patent No. 6083716
: GENERAL INFORMATION:
: APPLICANT: Wilson, James M.
: APPLICANT: Farina, Steven F.
: APPLICANT: Fisher, Krishna J.
: TITLE OF INVENTION: Chimpanzee Adenovirus Vectors
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Howson and Howson
: STREET: Spring House Corporate Cntr., P.O. Box 457
: CITY: Spring House
: STATE: Pennsylvania
: COUNTRY: United States of America
: ZIP: 19477
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/923,137
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/024,700
: FILING DATE: 06-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Bak, Mary E.
```

		Query Match	6.5%	Score 32.8	DB 4	Length 2613;
		Best Local Similarity	58.0%	Pred. No. 2	1.	
Matches	58; Conservative	0;	Mismatches	42;	Indels	0;
Gaps	0;					
QY	402	TTCCTACGACTTCTCTATAGCAGGTGAACAAGCTGCACATATTTTTAATCTCACT	461			
Dd	325	TTTACACAACTTTCTCCAGAATCAGATTGGATTAATATCTCTCAAAAATATTTTGTAACHCCT	266			
OY	462	GGCATATTCAAAGTTTCATTTGCTGTCTTAACAAGSGTGT	501			


```

? TITLE OF INVENTION: FUSIONS
? FILE REFERENCE: 00786/350007
? CURRENT APPLICATION NUMBER: US/09/244,796
? CURRENT FILING DATE: 1999-02-05
? EARLIER APPLICATION NUMBER: 60/035,963
? EARLIER FILING DATE: 1997-01-27
? EARLIER APPLICATION NUMBER: 60/064,491
? EARLIER FILING DATE: 1997-11-06
? EARLIER APPLICATION NUMBER: 09/007,005
? EARLIER FILING DATE: 1998-01-14
? NUMBER OF SEQ ID NOS: 33
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 17
? LENGTH: 289
? TYPE: RNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Translation template
? FEATURE:
? NAME/KEY: misc-feature
? LOCATION: (1)...(289)
? OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

```

```

Query Match Similarity      6.5%, Score 32.4; DB 4; Length 289;
Best Local Similarity      5.4%, Pred. No. 1.1;
Matches 11; Conservative   87; Mismatches 106; Indels    0; Gaps 0;

QY      191 TCACGCCCATCAGTGGGACCTCCACAGCACCAATGTTCAACGGATGACCACTGCTCTT 250
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      44 UIRACRARAARUGRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 103
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      251 CGGTGTGTGATTTACATCTCATGGCCCGGTGTGGGGACTTCTTTCATTTGGAAC 310
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      104 SNRRRSRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 163
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      311 CAGGATGCTTTTCAAAGCCCAATCACTGGGAGACCGAGACAGGAGAGACCAAGGGGAA 370
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      164 SNRRRSRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 223
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      371 GGGAGAGAAGCAATTAAGAACCA 394
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      224 GRCRURGRCGRURARARCRURCR 247
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-234-186-1/c
; Sequence 1, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
; APPLICANT: Horwitz, H. Robert
; APPLICANT: Hensalter, Michael
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 01997/201005
; CURRENT APPLICATION NUMBER: US/09/234,186
; CURRENT FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 07/998,933
; EARLIER FILING DATE: 1992-06-12
; EARLIER APPLICATION NUMBER: 07/927,681
; EARLIER FILING DATE: 1992-08-10
; EARLIER APPLICATION NUMBER: 08/288,295
; EARLIER FILING DATE: 1994-08-10
; EARLIER APPLICATION NUMBER: 08/801,248
; EARLIER FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6559
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)..(6559)
; OTHER INFORMATION: n = A,T,C or G
US-09-234-186-1

```

Query Match	6.5%;	Score 32.4;	DB 4;	length 6559;
Best Local Similarity	52.2%;	Pred. No. 4.3;		
Matches 72;	Conservative 0;	Mismatches 66;	Indels 0;	Gaps 0.

Oy	358	GACCAAGGGGAGGGGAGAGAAAGAAATTAAGAACACGTTATTTCTTAACAGACTTTC	417
Db	6430	GTCACATGAAAAAGCAGCCTAAGCTAGCTGTAAGCCAAACAATTTTATTAAGCTACATATCTTAA	6371
Oy	418	TATAGGAGTTGTGAAGAGTGTCACATATTTTAAATCTACACTGGCAATATTCAAAAGTT	477
Db	6370	AAATGTATTTGTTTAAAGGAACACGTAGTTTCTGAGTGGGCTTGGCCACGATCACAGTTC	6311
Oy	478	TTCAATTGTCCTAACAA	495
Db	6310	ATAGTGGTCCTGCAGAA	6293

```

RESULT 13
US-09-233-527-1/c
: Sequence 1, Application US/09233527
: Patent No. 6465617
: GENERAL INFORMATION:
: APPLICANT: Horvitz, H. Robert
: APPLICANT: Hengartner, Michael
: TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
: TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED AND
: TITLE OF INVENTION: USES THEREFOR
: FILE REFERENCE: 01997/201004
: CURRENT APPLICATION NUMBER: US/09/233.527
: PRIOR FILING DATE: 1999-01-20
: PRIOR APPLICATION NUMBER: 07/898,933
: PRIOR FILING DATE: 1992-06-12
: PRIOR APPLICATION NUMBER: 07/927,681
: PRIOR FILING DATE: 1992-08-10
: PRIOR APPLICATION NUMBER: 08/288,295
: PRIOR FILING DATE: 1994-08-10
: PRIOR APPLICATION NUMBER: 08/801,248
: PRIOR FILING DATE: 1997-02-19
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 6559
: TYPE: DNA
: ORGANISM: Caenorhabditis elegans
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(6559)
: OTHER INFORMATION: n = A,T,C or G
US-09-233-527-1

```

	Query Match	6.5%;	Score 32.4;	DB 4;	Length 6559;
	Best Local Similarity	52.2%;	Pred. No. 4.3;	Mismatches 66;	Indels 0; Gaps 0;
	Matches	72;	Conservative	0;	
QY	358	GACCAAGGGAAGGGAGAANAAGAAATAAAGACAACGTTATTTCTTAACAGACTTTC			
Dd	6430	GTCGACTGAAAAAGCAGCCCTAGACTGCTAGGCCAACAAATTTTATNCCACAAATCTTAA			
QY	418	TATAGSAGTTGTAAAGAAGTGACATATTTTTTAAATCTCAGCGCAATATTCAAAGTT			
Dd	6370	AATGATATTTGTTTAAAGSAAACAGTAGTTTCTGATGGGCTTCGCACGATTCACAGTC			
QY	478	TTCATTGTGCTTTACAA			
Dd	6310	ATAGTGGTGGCTGCAGAA			

RESULT 14


```

PCT-US93-05651-1/c
Sequence 1, Application PC/TUS9305651
GENERAL INFORMATION:
TITLE OF INVENTION: A Gene Which Prevents Programmed Cell Death
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05651
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-05651-1

Query Match 6.5%; Score 32.4; DB 5; Length 6560;
Best Local Similarity 52.2%; Pred. No. 4.3;
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0

QY 358 GACCAAGGGAGGAGGAGCAAGAAATTAAGAACACGTTATTCTTAACAGACTTTC 417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6431 GTCGACTGAAAAAGGAGCGCTAGCTGCTGACCAACATTTTTATACCTACATACTTAA 6372
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 418 TATAGAGATTGTATAGAGGTGCACATATTTTAAATCTCAGCGGAATATCAAGTT 477
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6371 AAATGATATTGTTTAAGGAACACGTAAGTTTTCGTAGTGGGTCTTGCCACAGACAGTC 6312
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 478 TTCATTGTGCTTAACAA 495
      | | | | | | | | | |
Db 6311 ATAGTGTGCTGCACAGA 6294
      | | | | | | | | | |

RESULT 15
US-08-471-033-39
Sequence 39, Application US/08471033
Patent No. 5770696
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Kozief, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,033
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994

```

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/037,057
? FILING DATE: 25-MAR-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Pace, Gary M.
? REGISTRATION NUMBER: P-40,403
? REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-541-8582
? TELEFAX: 919-541-8689
? INFORMATION FOR SEQ ID NO: 39:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1241 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: other nucleic acid
? DESCRIPTION: /desc = "Synthetic DNA"
? HYPOTHETICAL: NO
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 9..1238
? OTHER INFORMATION: /note= "Maize optimized DNA
? OTHER INFORMATION: sequence encoding VIP2A(a) with the Bacillus secretion sig
? OTHER INFORMATION: removed as contained in PCIB5327"
US-08-471-033-39

Query Match          6.4%; Score 32; DB 1; Length 1241;
Best Local Similarity 51.4%; Pred. No. 2.7;
Matches 74; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

OY      59   ACCAGCTGGAGTATTCAGGAAGAATGAAAGCCACTNAGGGAATGGCGAAGACGTTT 118
           || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      139   ACTTCCTCGAACACAGACAGACATCAAGACCACTCAAGAGCATCCCTTCAGCATAG 198
OY      119   CTGAATCATGCATGAGCAGATCTGCCCTCTGGAGGAGAAAGCAGCGCTTTACCAGATT 178
           || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      199   CCGGAGCTGTGAGAGACGAGATCAAGGACCTGAGAGGAGATGCACAAGATGTTCCACANA 258
OY      179   CCCTTCACATCTTCACAGCCATCA 202
           || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      259   CCAACCTGAGCACACAGCATCATCA 282

```

Search completed: February 7, 2003, 09:06:22
Job time : 63.6858 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:01:14 ; Search time 32.3549 Seconds
(without alignments)
7339.716 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300
 Perfect Name: E01

Sequence: 1 caattgtgtgaagcttcg.....tgtgtccttaacaagtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

1: /cgn2.6/ptodata/2/pubpna/us07_PUBCOMB.seq: *

2: /cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *

3: /cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq: *

4: /cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq: *

5: /cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq: *

6: /cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *

7: /cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq: *

8: /cgn2.6/ptodata/2/pubpna/US08_PUBCOMB.seq: *

9: /cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq: *

10: /cgn2.6/ptodata/2/pubpna/US10_PUBCOMB.seq: *

11: /cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq: *

12: /cgn2.6/ptodata/2/pubpna/US10_PUBCOMB.seq: *

13: /cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq: *

14: /cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	336	67.1	487	10	US-09-998-5598-1612		Sequence 1612, Ap
2	139	27.7	1031	10	US-09-876-889-157		Sequence 157, Ap
3	71	14.2	423	10	US-09-960-352-3043		Sequence 3043, Ap
4	67.4	13.5	4391	10	US-09-738-969A-7		Sequence 7, Ap
5	67.4	13.5	6454	10	US-09-738-969A-7		Sequence 11, Ap
6	50	10.0	434	9	US-09-796-692-4813		Sequence 4813, Ap
7	44	8.8	505	9	US-09-796-692-5039		Sequence 5039, Ap
8	44	8.8	505	9	US-09-796-692-5523		Sequence 5223, Ap
9	41.8	8.3	244	10	US-09-738-960-84		Sequence 84, Ap
10	41.8	8.3	4026	10	US-09-738-960-3		Sequence 3, Ap
11	41.8	8.3	7215	10	US-09-738-960-1		Sequence 1, Ap
12	41.8	8.3	66686	10	US-09-738-960-86		Sequence 86, Ap
13	37.2	7.4	521	10	US-09-964-824A-1		Sequence 1, Ap
14	37.2	7.4	521	10	US-09-954-456-1193		Sequence 1193, Ap
15	36.6	7.3	696	10	US-09-815-242-7900		Sequence 7900, Ap
16	36.4	7.3	143068	10	US-09-967-768A-316		Sequence 316, Ap
17	35.2	7.0	126512	10	US-09-804-474A-3		Sequence 3, Ap
18	35	7.0	23626	10	US-09-764-878-261		Sequence 261, Ap
19	35	7.0	23626	10	US-09-764-860-940		Sequence 940, Ap

20	35	7.0	23632	10	US-09-764-868-261	Sequence 262, Appl
21	35	7.0	23632	10	US-09-764-868-941	Sequence 941, Appl
22	34.8	6.9	429	10	US-09-800-729-150	Sequence 150, Appl
23	34.8	6.9	650	10	US-09-815-242-7864	Sequence 7864, Appl
24	34.8	6.9	368004	10	US-09-949-654-3	Sequence 3, Appl
25	32.8	6.5	1461	10	US-09-880-192-742	Sequence 42, Appl
26	32.8	6.5	3891	12	US-10-051-952-1	Sequence 1, Appl
27	32.8	6.5	4835	10	US-09-288-326-10	Sequence 10, Appl
28	32.6	6.5	245	10	US-09-864-761-23212	Sequence 23212, Appl
29	32.6	6.5	469	10	US-09-864-761-6499	Sequence 6499, Appl
30	32.6	6.5	563	10	US-09-833-381-542	Sequence 542, Appl
31	32.6	6.5	4753	9	US-10-098-841-18	Sequence 18, Appl
32	32.4	6.5	695	9	US-10-016-157A-6	Sequence 69, Appl
33	32	6.4	1767	10	US-09-864-761-60592	Sequence 30592, Appl
34	32	6.4	2604	10	US-09-529-063-9	Sequence 9, Appl
35	31.8	6.3	384	10	US-09-983-965-198	Sequence 198, Appl
36	31.8	6.3	1750	10	US-09-925-301-633	Sequence 633, Appl
37	31.6	6.3	337	10	US-09-864-761-22871	Sequence 22871, Appl
38	31.4	6.3	740	10	US-09-733-705-6	Sequence 6, Appl
39	31.4	6.3	740	10	US-09-850-716A-6	Sequence 6, Appl
40	31.4	6.3	740	10	US-09-891-778-6	Sequence 6, Appl
41	31.4	6.3	1249	10	US-09-925-301-501	Sequence 501, Appl
42	31.4	6.3	4218	10	US-09-755-639-8	Sequence 8, Appl
43	31.4	6.3	4218	10	US-09-984-198-8	Sequence 8, Appl
44	31.4	6.3	202001	10	US-09-734-674-3	Sequence 3, Appl
45	31.2	6.2	2000	9	US-09-938-842A-9943	Sequence 4943, Appl

ALIGNMENTS

```

RESULT 1
US-09-998-598-1612
: Sequence 1612, Application US/09988598
: Patent No. US20020150922A1
: GENERAL INFORMATION:
: APPLICANT: Stolk, Johan A.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Chenault, Ruth A.
: APPLICANT: Meagher, Madelaine Joy
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: TITRE OF INVENTION: DIAGNOSIS OF COLON CANCER
: FILE REFERENCE: 210121.561
: CURRENT APPLICATION NUMBER: US/09/998,598
: CURRENT FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 2606
: SOFTWARE: Corixa Invention Disclosure Database
: SEQ ID NO 1612
: LENGTH: 487
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-998-598-1612

```

```
Query Match      67.18; Score 336; DB 10; Length 487;
Best Local Similarity -100.0%; Pred. No. 2.2e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY	1	CAATTGTGSAACCTTGGCGCTAACCCCTTACGGTAAACAAGCTGTGATTAAACAAGC	60
Db	152	CAATTGTGSAACCTTGGCGCTAACCCCTTACGGTAAACAAGCTGTGATTAAACAAGC	21
OY	61	CAGCTCAGATTCAGSANAAMATGAACCCAACTCAGGAAATGGCGAAGAGCTTCT	120
Db	212	CAGCTCAGATTCAGSANAAMATGAACCCAACTCAGGAAATGGCGAAGAGCTTCT	27
OY	121	GAAATCATGTCATGAGCAGATCTGGCCCCCTGGAGSAGSAGCAGCGCTCTTACCGAATTC	180
Db	272	GAAATCATGTCATGAGCAGATCTGGCCCCCTGGAGSAGSAGCAGCGCTCTTACCGAATTC	333
OY	181	CTTCACATCTTCACGCCCATCAGTGGGACCTCAACAAGCAAGCAATGGTTACAGGGATGAGC	240
Db	332	CTTCACATCTTCACGCCCATCAGTGGGACCTCAACAAGCAAGCAATGGTTACAGGGATGAGC	391

```

/0.48; Pred: NO. 8.9e-12;
Mismatches 40: Indels 0: Gaps 0:

```

67.48; Pred. No. 3.5e-10;

	Matches	95; Conservative	0;	Mismatches	46; Indels	0; Gaps	0;
QY	2	AATTGTGGAAGCTTCGGTCAAGCCTTAACCGTAACAACTGCTGATTTAAAGAAGACC	61				
Db	3802	AATTTATACAGACATGCACATTTGCCATTGCACCTTAACCTAATTAAGACGCGCTAATTAAGAAGATC	3861				
QY	62	AGCTCGNGTATCAGGAAGAATAAGAACCAACTACAGGGAAATGGCCAGSAGCTTCTG	121				
Db	3862	AAGTTTAGTACCTGTAAGGGCTTAAAGTCAAATTTCCAGAGCATGTGTAAGAATTTATCTG	3921				
QY	122	AAATCATGCATGAGCAATCT	142				
Db	3922	ACAATTATCCATGACAGATAT	3942				

```

RESULT 5
US-09-736-969A-1
; Sequence 1, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736,969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6454
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length human CLASP-4 cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(6121)
; OTHER INFORMATION: human CLASP-4
US-09-736-969A-1

```

	Best Local Similarity	67.4%	Pred. No.	4.3e-10;	Mismatches	95; Conservative	0; Mismatches	46; Indels	0; Gaps	0;
OY	2	AATTTGTGGAAGCTTSCGGTCACAGCCTTAAGCGGTAAACGACACTGTGATTAAAGAAGACC	61							
Dd	5865	AATTTAATACACAGCATGCACCAATTGCACTTGAACTAAATATGACGGCTATAATTTAAAAGAAGATC	5924							
OY	62	AGCTCGAGTATCAGAGAACATGAAGAAGCAACTACAGGGAAATGCCAGAGACTTCTTG	121							
Dd	5925	AAGTGGAGTACCATGAAGGGCTAAAGTCMAATTCACAGACATGTAAAGAAATTATCTTG	5984							
OY	122	AAATCATGCATGACGACAATCT	142							
Dd	5985	ACATTATCATCTGAGCAGATAT	6005							

```

RESULT 6
US-09-796-692-4813
; Sequence 4813, Application US/09796692
; Publication NO. US20020198362A1
;
GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077_001200
;
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
;
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
;
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
;
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
;
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
;
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
;
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
;
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
;
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
;
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
;
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
;
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
;
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
;
NUMBER OF SEQ ID NOS: 9597
;
SOFTWARE: FastSeq for Windows Version 3.0
;
SEQ ID NO 4813
;
LENGTH: 434
;
TYPE: DNA
;
ORGANISM: Homo sapiens
;
US-09-796-692-4813

Query Match      10.0%; Score 50; DB 9; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 452 AATATCTACGTGCATATTTCAAGTTTTCATGTGCTTTACCAAGGTGT 501
DB 1 AATCTCACTGCGCATATTTCAAGTTTTCATGTGCTTTACCAAGGTGT 50

RESULT 7
US-09-796-692-5039
; Sequence 5039, Application US/09796692

```

```
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5039
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (341)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-5039

Query Match      8.8%; Score 44; DB 9; Length 505;
Best Local Similarity 65.0%; Pred. No. 0.0013;
Matches 65; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 64 CTCGAGTATCAGGAAGAAATGAAAGCCCACTACAGGGAATGCGGAGAGACTTTCTGAA 123
DB 1 CTGGAGTACAGGAGAACTGAGGTCCACTACAGGAGCACTGCTCAGCACTCCACA 60
QY 124 ATCATGATGACGAGATCTGCCCTCGAGGAGAGAGACA 163
DB 61 GTCATGAATGACGACATTACGGGCGAGGAGCAGCTGTCAA 100

RESULT 8
US-09-796-692-5223
; Sequence 5223, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
```

```
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5223
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-5223

Query Match      8.8%; Score 44; DB 9; Length 505;
Best Local Similarity 65.0%; Pred. No. 0.0013;
Matches 65; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 64 CTCGAGTATCAGGAAGAAATGAAAGCCCACTACAGGGAATGCGGAGAGACTTTCTGAA 123
DB 1 CTGGAGTACAGGAGAACTGAGGTCCACTACAGGAGCACTGCTCAGCACTCCACA 60
QY 124 ATCATGATGACGAGATCTGCCCTCGAGGAGAGAGACA 163
DB 61 GTCATGAATGACGACATTACGGGCGAGGAGCAGCTGTCAA 100

RESULT 9
US-09-736-960-84
; Sequence 84, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia, III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-00051105
; CURRENT APPLICATION NUMBER: US/09/736,960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
```

```
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 21st exon
US-09-736-960-84
```

```
Query Match      8.3%; Score 41.8; DB 10; Length 244;
Best Local Similarity 56.0%; Pred. No. 0.004;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
QY 16 TGGCGTCAACCTTACGCGTAACGACGCTGTGATTAAAGAACCCAGCTCGAGTTCAG 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 TGTGTGAACCTGTAAGAAAACAGCGCTCTCATCAGCGCAGACGAGGGAATTCAG 97

QY 76 GAAGAAATGAAGCCACACTACAGGGAATGGCGAAGAGAGCTTCTGAATTCATGTCATGAG 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 CAGGAACCTCAAAAAGAACTTATACAGCTTAAGAGAACTCAGGCCATGATCGAGCGG 157

QY 136 CAGATCTGCCCTCGAGAGAG 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 158 AAATTCAGAACTGTACAAG 178
```

```
RESULT 10
US-09-736-960-3
; Sequence 3, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
```

```
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4026
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: preliminary human CLASP-5 CDNA sequence
; NAME/KEY: CDS
; LOCATION: (1)..(3066)
; OTHER INFORMATION: human CLASP-5
US-09-736-960-3
```

```
Query Match      8.3%; Score 41.8; DB 10; Length 4026;
Best Local Similarity 56.0%; Pred. No. 0.017;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
QY 16 TGGCGTCAACCTTACGCGTAACGACGCTGTGATTAAAGAACCCAGCTCGAGTTCAG 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2740 TGTGTGAACCTGTAAGAAAACAGCGCTCTCATCAGCGCAGACGAGGGAATTCAG 2799

QY 76 GAAGAAATGAAGCCACACTACAGGGAATGGCGAAGAGAGCTTCTGAATTCATGTCATGAG 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2800 CAGGAACCTCAAAAAGAACTTATACAGCTTAAGAGAACTCAGGCCATGATCGAGCGG 2859

QY 136 CAGATCTGCCCTCGAGAGAG 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2860 AAATTCAGAACTGTACAAG 2880
```

```
RESULT 11
US-09-736-960-1
; Sequence 1, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
```

```

; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length human CLASP-5 cDNA
; NAME/KEY: CDS
; LOCATION: (112)..(6159)
; OTHER INFORMATION: human CLASP-5
US-09-736-960-1
```

```

Query Match      8.3%; Score 41.8; DB 10; Length 7215;
Best Local Similarity 56.0%; Pred. No. 0.022; Mismatches 62; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```

QY 16 TCGGCTCAACCTTACCGCTTAACGACGCTGTGATTAAAGAACCCAGCTCGAGTATCAG 75
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5929 TGTGTGAAGCTGTAGAGAAACAGCGCTCATCAGCGAGCCAGAGGGAATATCAG 5988
QY 76 GAGGAATGAAGCCCAACTACAGGGAATGCGCAAGGACCTTTCTAATCATGCTGAG 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5989 CAGGAACCTCAAAAGAACTATACAGCTAAAGAGAACCTCAGGCCAATGATCGAGCG 6048
QY 136 CAGATCTGCCCTCGAGAGG 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6049 AAAATTCCAGAACTGTACAAG 6069
```

```

RESULT 12
US-09-736-960-86
; Sequence 86, Application US/09736960
; Patent No. US2002010267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
```

```

; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 6686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ordered human genomic DNA at CLASP-5 locus
US-09-736-960-86
```

```

Query Match      8.3%; Score 41.8; DB 10; Length 6686;
Best Local Similarity 56.0%; Pred. No. 0.068; Mismatches 62; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```

QY 16 TCGGCTCAACCTTACCGCTTAACGACGCTGTGATTAAAGAACCCAGCTCGAGTATCAG 75
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62399 TGTGTGAAGCTGTAGAGAAACAGCGCTCATCAGCGAGCCAGAGGGAATATCAG 62458
QY 76 GAGGAATGAAGCCCAACTACAGGGAATGCGCAAGGACCTTTCTAATCATGCTGAG 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62459 CAGGAACCTCAAAAGAACTATACAGCTAAAGAGAACCTCAGGCCAATGATGAGCGG 62518
QY 136 CAGATCTGCCCTCGAGAGG 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62519 AAAATTCCAGAACTGTACAAG 62539
```

```

RESULT 13
US-09-964-824A-1
; Sequence 1, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(521)
; OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-1
```

```

Query Match      7.4%; Score 37.2; DB 10; Length 521;
Best Local Similarity 57.4%; Pred. No. 0.14; Mismatches 49; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```

QY 360 CCAAGGGGAAGGCGAGAGAAAGAAATTAAGAACACGCTATTCTTAACAGACTTTCTA 419
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 245 CCGGGGGTTGGGGAGGGGAGTCGATTAATTAATAAAAGTTTAGAAGGCCATGAGTAAATA 304

QY 420 TAGAGTTTGAAGAGTGCACATATTTTTTAAATCTCAGCGCAATTTCAA 474

Db 305 TCGAATTAAGTATGAAATTTAATATTAATCTTTTAAAGGGGTAGGCAATGATGAAA 359

RESULT 14
US-09-954-456-1193
; Sequence 1193, Application US/09954456
; Patent No. US2002115057A1
; Summary: Tumor necrosis factor

```

APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer
TITLE OF INVENTION: Sels
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIORITY APPLICATION NUMBER: US/60/233,617
PRIORITY FILING DATE: 2000-09-18
PRIORITY APPLICATION NUMBER: US/60/234,052
PRIORITY FILING DATE: 2000-09-20
PRIORITY APPLICATION NUMBER: US/60/234,923
PRIORITY FILING DATE: 2000-09-25
PRIORITY APPLICATION NUMBER: US/60/235,134
PRIORITY FILING DATE: 2000-09-25
PRIORITY APPLICATION NUMBER: US/60/235,637
PRIORITY FILING DATE: 2000-09-26
PRIORITY APPLICATION NUMBER: US/60/235,638
PRIORITY FILING DATE: 2000-09-26
PRIORITY APPLICATION NUMBER: US/60/235,711
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: US/60/235,720
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: US/60/235,840
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: US/60/235,863
PRIORITY FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1193
LENGTH: 521
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1193

```

Query Match	7.48;	Score 37.2;	DB 10;	Length 521;
Best Local Similarity	57.48;	Pred. No. 0.14;		
Matches 66;	Conservative 0;	Mismatches 49;	Indels 0;	Gaps 0

QY	360	CCAAAGGCGAAGGGCAGAGAAAGAAAGAAAGAACACGCTATTTCCTAACAGACTTCCTA	419
Db	245	CCGGGGCGTTGGGGGAGGAGTGGATTAAATAAAAGTTTAAAGCCCATGNNATTAATA	304
QY	420	TGAGAGTTGTAAAGAGGCGCACATATTTTAAATCTCAGTGGCAATATTCAAA	474
Db	305	TGCAATATGTATGAATTTTAAATATATCTTTTAAAGGGGTGTGGCAAAATGAATGAA	359

RESULT 15
US-09-815-242-7900
: Sequence 7900, Application US/09815242
: Patent No. US20020061569A1

: APPLICANT: Haselbeck, Robert
 : APPLICANT: Ohlsen, Kari L.
 : APPLICANT: Zyskind, Judith W.
 : APPLICANT: Wall, Daniel
 : APPLICANT: Trawick, John D.
 : APPLICANT: Carr, Grant J.

```

? APPLICANT: Yamamoto, Robert T.
? APPLICANT: Xu, H. Howard
? TITLE OF INVENTION: Identification of Essential Genes in
? TITLE OF INVENTION: Prokaryotes
? FILE REFERENCE: ELITRA.011A
? CURRENT APPLICATION NUMBER: US/09/845,242
? CURRENT FILING DATE: 2001-03-21
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? NUMBER OF SEQ. ID NOS: 14110
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 7900
? LENGTH: 696
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(696)
US-09-815-242-7900

```

Query Match	7.3%	Score 36.6;	DB 10;	Length 696;
Best Local Similarity	52.3%;	Pred. No. 0.25;		
Matches 81; Conservative	0;	Mismatches 74;	Indels 0;	Gaps 0;

QY 128 TGGATGAGCAGATCTTGCCCCCTGGGAGGGAAGACGAGCGCTTACGAGTAATTCCTTCACA 187

Db 249 TGGCGAAGCCGCTCTGCGCTGCGCGCGCCGACGAAGTGGGTATGAGCAGAGCTGGCTGCCGA 308

QY 188 TCTTCACGCCATCATGATGGGAGCTCCACACAGCAGCATATGGTTCACGGAGTGAACAGACTCTGT 247

Db 309 AATGAAAGGGGGCGACACTGAACACTGACAGCTGTCATCTTCCCGAGATGCCATGCTGT 368

QY 248 CTTTCGCTGCTGATTTACATCTCATATGGCCGCTGTG 282

Db 369 CGTGGTTCAGCTGGGCCAGATCTCTGGGCGCCGCG 403

```
Search completed: February 7, 2003, 09:15:32
Job time : 83.3549 secs
```

